

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

KARMA VERİ MODELLERİNİN OPTİMİZASYONU İÇİN ESNEK
HESAPLAMA YÖNTEMLERİ

Gözde KARAKOÇ

İSTATİSTİK ANABİLİM DALI

ANKARA
2024

Her hakkı saklıdır

ÖZET

Doktora Tezi

KARMA VERİ MODELLERİNİN OPTİMİZASYONU İÇİN ESNEK HESAPLAMA YÖNTEMLERİ

Gözde KARAKOÇ

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Özlem TÜRKŞEN

Gerçek dünya problemlerine ilişkin verilerin hem nitel hem de nicel değerli olması durumunda problem karma veri problemi olarak adlandırılır. Bu tez çalışmasında, karma veri probleminin modellenmesi ve çok amaçlı optimizasyonu amaçlanmıştır. Çok yanıtlı karma veri probleminin modelleme aşamasında, Genelleştirilmiş Doğrusal Modeller (Generalized Linear Models-GLM), Ridge, LASSO, Elastik Net ve Görünüşte İlişkisiz Regresyon (Seemingly Unrelated Regression-SUR) analizleri yapılarak tahmini yanıt fonksiyonları elde edilmiştir. Modelleme aşamasında elde edilen tahmini yanıt fonksiyonları amaç fonksiyonları olarak ele alınıp karma verili çok amaçlı optimizasyon (Multi-Objective Optimization - MOO) problemi oluşturulmuştur. Çalışmada, bu tür problemlerin Pareto çözüm kümesinin bulunması için popülasyon tabanlı bir yapay zeka optimizasyon algoritması olan Baskın Sıralı Genetik Algoritma-II (Nondominated Sorting Genetic Algorithm-II / NSGA-II) ve Baskın Sıralı Genetik Algoritma-III (NSGA-III) uyarlanmıştır. NSGA-II'nin ve NSGA-III'ün değişken gösterimi, başlangıç popülasyonu oluşturma ve genetik operatörlerin uygulanması aşamalarında yapılan uyarlamalarla elde edilen yöntemler Karma Verili Baskın Sıralı Genetik Algoritma-II (Mixed Data Nondominated Sorting Genetic Algorithm-II / MDNSGA-II) ve Karma Verili Baskın Sıralı Genetik Algoritma-III (MDNSGA-III) olarak adlandırılmıştır. Önerilen MDNSGA-II ve MDNSGA-III yöntemleriyle kesikli girdi değişkenlerinin tanım kümesinden değerler alması sağlanarak karma verili çok amaçlı problem optimize edilmiştir. Çalışmanın uygulama aşamasında, UCI Machine Learning Repository veri tabanındaki enerji verimliliği karma veri kümesine ve gıda alanında literatürde mevcut olan deneysel veri kümesinin karma verili hale getirilmesiyle elde edilen veri kümesine MDNSGA-II ve MDNSGA-III yöntemleri uygulanmıştır. MOO sonucunda, MDNSGA-III ile elde edilen Pareto çözüm kümesinde MDNSGA-II'ye göre iyileşme sağlandığı ve daha fazla çözüm çeşitliliğinin elde edilebildiği sonucuna ulaşılmıştır.

Temmuz 2024, 149 sayfa

Anahtar Kelimeler: Karma veri, GLM, Ridge, LASSO, Elastik Net, SUR, esnek hesaplama yöntemleri, çok amaçlı optimizasyon problemi, NSGA-II, NSGA-III, MDNSGA-II, MDNSGA-III

ABSTRACT

PhD Thesis

SOFT COMPUTING METHODS FOR OPTIMIZATION OF MIXED DATA MODELS

Gözde KARAKOÇ

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Prof. Dr. Özlem TÜRKŞEN

The problem is called mixed data problem, if the data has both qualitative and quantitative values for a real world problem. In this thesis, modeling and multiobjective optimization of the mixed data problem is aimed. In the modeling phase of the multi response mixed data problem, Generalized Linear Models (GLM), Ridge, LASSO, Elastic Net and Seemingly Unrelated Regression (SUR) analyses are performed to obtain estimated response functions. The estimated response functions obtained in the modeling phase are taken as objective functions and a mixed data multiobjective optimization problem (MOO) is formed. In this study, Nondominated Sorting Genetic Algorithm-II (NSGA-II) and Nondominated Sorting Genetic Algorithm-III (NSGA-III), which are population based artificial intelligence optimization algorithms, are adapted to find the Pareto solution set of such problems. The methods obtained by adapting NSGA-II and NSGA-III in the variable representation, initial population generation and application of genetic operators are named Mixed Data Nondominated Sorting Genetic Algorithm-II (MDNSGA-II) and Mixed Data Nondominated Sorting Genetic Algorithm-III (MDNSGA-III). With the proposed MDNSGA-II and MDNSGA-III methods, the mixed data multiobjective problem is optimized by ensuring that the discrete input variables take values from the definition set. In the application phase of the study, the MDNSGA-II and the MDNSGA-III methods are applied to the energy efficiency mixed dataset from the UCI Machine Learning Repository database and the dataset obtained by converting the experimental dataset available in the literature in the food field into mixed data. As a result of the MOO, it is concluded that the Pareto solution set obtained with the MDNSGA-III, have more solution diversity with improvement than the Pareto solution set of the MDNSGA-II.

July 2024, 149 page

Key Words: Mixed data, GLM, Ridge, LASSO, Elastic Net, SUR, soft computing methods, multiobjective optimization problem, NSGA-II, NSGA-III, MDNSGA-II, MDNSGA-III

TEŞEKKÜR

Çalışmalarımı yönlendiren, araştırmalarımın her aşamasında bilgi, öneri, tecrübe ve yardımlarını esirgemeyerek bana her zaman yol gösteren ve çalışma azmini kendime örnek aldığım değerli danışman hocam Sayın Prof. Dr. Özlem TÜRKŞEN (Ankara Üniversitesi, İstatistik A.B.D.)'e en içten saygı ve minnetlerimi sunarım.

Değerli TİK üyesi hocalarım, Sayın Prof. Dr. Ayşen APAYDIN (Ankara Üniversitesi, Aktüerya Bilimleri A.B.D.)'a, Sayın Prof. Dr. Halil AYDOĞDU (Ankara Üniversitesi, İstatistik A.B.D.)'ya ve bana lisans dönemi öğrenim sürecinde Optimizasyon dersini sevdiren ve doktora tezimi bu alanda yapmama vesile olan Sayın Prof. Dr. Hacı Hasan ÖRKÜ (Gazi Üniversitesi, İstatistik A.B.D.)'ye olumlu eleştirileri ve desteklerinden dolayı teşekkürlerimi bir borç bilirim.

Ayrıca Ankara Üniversitesi İstatistik ve Matematik Bölümü öğretim üyelerine, desteğini her zaman hissettiğim ve bu süreçte beni yüreklendiren iş arkadaşlarıma, bilgisayar yazılım programları konusundaki bilgilerini benimle paylaşan ve desteğini hiçbir zaman esirgemeyen arkadaşım Arş. Gör. Dr. Emre KOÇAK (Gazi Üniversitesi, İstatistik A.B.D.)'a teşekkürlerimi sunarım.

Bana her konuda destek olan ve her zaman inanan, gücümü ve azmimi bana duyduğu güven ve sevgisinden aldığım, beni bugünlere getiren ve hakkını hiçbir zaman ödeyemeyeceğim fedakar annem Şenay ÖZÇİRPAN'a, desteklerini hiçbir zaman esirgemeyen canımdan öte kardeşlerim Mertcan ÖZÇİRPAN ile Özgü KARAGÜLER ve ailelerine teşekkürlerimi sunarım. Ayrıca beni kendi evlatlarından ayırmayan ve her zaman desteklerini hissettiğim eşimin ailesine teşekkürü bir borç bilirim. Bu süreçte bir an bile yanımdan ayrılmayan, koşulsuz sevgisini ve her konudaki desteğini hep hissettiren, hayata karşı duruşuyla her zaman gurur duyduğum, eşim Serhat KARAKOÇ'a katlandığı tüm zorluklar ve bana yaşattığı tüm güzellikler için tüm kalbimle ve en derin duygularıyla teşekkür ederim. Hayat enerjim olan, enerjisiyle girdiği tüm ortamı aydınlatan ve bana varlığıyla her zaman güç veren kızım Ada KARAKOÇ, Ulu Önder Mustafa Kemal ATATÜRK'ün izinde, çağdaş bilimin ışığında hep parla. Yaşamım boyunca her zaman yanımda olan, verdiği hayat mücadelesiyle bana her daim örnek olan, evlatları için koşulsuz her zorluğu göğüsleyen, maddi ve manevi desteğini hiçbir zaman esirgemeyen 2020 yılında kaybettiğim ama yaşadığım süre boyunca her zaman kalbimde yaşatacağım çınarım, babam, Hüseyin ÖZÇİRPAN'ın anısına...

Gözde KARAKOÇ
Ankara, Temmuz 2024

İÇİNDEKİLER

TEZ ONAYI

ETİK.....	i
ÖZET.....	ii
ABSTRACT	iii
TEŞEKKÜR	iv
SİMGELER ve KISALTMALAR DİZİNİ	ix
ŞEKİLLER DİZİNİ	xii
ÇİZELGELER DİZİNİ	xiv
1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR	1
1.1 Giriş	1
1.2 Önceki Çalışmalar	5
2. İSTATİSTİKSEL MODELLEME	12
2.1 Genelleştirilmiş Doğrusal Modeller	12
2.2 Görünüşte İlişkisiz Regresyon Modeli	18
2.3 İstatistiksel Daraltıcı Modeller	20
2.3.1 Ridge regresyon.....	21
2.3.2 LASSO regresyon.....	22
2.3.3 Elastik Net regresyonu	24
3. OPTİMİZASYON MODELLERİNİN SINIFLANDIRILMASI.....	27
3.1. Amaç Fonksiyonu Sayısına Göre Optimizasyon	28
3.1.1 Tek amaçlı optimizasyon	28
3.1.2 Çok amaçlı optimizasyon	29
3.2 Fonksiyon Yapılarına Göre Optimizasyon	33
3.2.1 Doğrusal optimizasyon	34
3.2.2 Doğrusal olmayan optimizasyon	34
3.3 Değişken Türlerine Göre Optimizasyon	35
3.3.1 Sürekli optimizasyon	35
3.3.2 Kesikli optimizasyon	35
3.3.3 Karma optimizasyon	38
4. OPTİMİZASYON PROBLEMLERİ İÇİN GENETİK ALGORİTMA	42
4.1 Genetik Algoritma.....	42

4.1.1 Karar değişkenlerinin gösterimi.....	44
4.1.2 Başlangıç popülasyonunun oluşturulması	45
4.1.3 Genetik operatörlerin uygulanması	46
4.2 Genetik Algoritmanın Teorik Temelleri	49
4.3 Tek Amaçlı Optimizasyon Problemleri için Genetik Algoritma	54
4.4 Çok Amaçlı Optimizasyon Problemleri için Genetik Algoritma	57
4.4.1 NSGA-II yöntemi	58
4.4.2 NSGA-III yöntemi.....	65
4.4.3 NSGA-II ile NSGA-III yöntemlerinin karşılaştırılması	79
5. KARMA OPTİMİZASYON PROBLEMLERİ İÇİN UYARLANMIŞ GENETİK ALGORİTMA	81
5.1 Uyarlanmış Genetik Algoritma.....	81
5.1.1 Uyarlanmış karar değişkenlerinin gösterimi.....	81
5.1.2 Uyarlanmış başlangıç popülasyonunun oluşturulması	82
5.1.3 Uyarlanmış genetik operatörlerin uygulanması	83
5.2 Tek Amaçlı Karma Optimizasyon Problemleri için Uyarlanmış Genetik Algoritma	86
5.3 Çok Amaçlı Karma Optimizasyon Problemleri için Uyarlanmış Genetik Algoritma	89
5.3.1 MDNSGA-II yöntemi	89
5.3.2 MDNSGA-III yöntemi	91
6. UYGULAMA.....	94
6.1 Uygulama 1 (Enerji Verimliliği Karma Veri Kümesi)	94
6.2 Uygulama 2 (Gıda Karma Veri Kümesi)	108
7. TARTIŞMA VE SONUÇ.....	119
KAYNAKLAR	121
EKLER.....	132
EK 1 Enerji karma verisi için $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi.....	133
EK 2 Enerji karma verisinde IRLS modeli için $\hat{Y}_1^{IRLS}$ ve $\hat{Y}_2^{IRLS}$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi	136
EK 3 Enerji karma verisinde IRLS modeli için $\hat{Y}_1^{IRLS}$ ve $\hat{Y}_2^{IRLS}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi	138

EK 4 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin	
MDNSGA-II ile elde edilen Pareto çözüm kümesi	140
EK 5 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin	
MDNSGA-III ile elde edilen Pareto çözüm kümesi	142
EK 6 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin	
MDNSGA-II ile elde edilen Pareto çözüm kümesi	144
EK 7 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-III	
ile elde edilen Pareto çözüm kümesi.....	146
ÖZGEÇMİŞ.....	148



SİMGELER DİZİNİ

α	[0, 1] aralığından seçilen rastgele bir sayı
$E(.)$	Beklenen değer
η	GLM’de doğrusal yanıt fonksiyonu
$exp(.)$	Üstel fonksiyon
$f(.)$	Olasılık yoğunluk fonksiyonu
$g(.)$	GLM’de bağ fonksiyonu
lb	Değişkenin aldığı alt sınır
$ln(.)$	Doğal logaritma fonksiyonu
μ	Konum parametresi
η_c	Çaprazlama indeksi
η_m	Mutasyon indeksi
n_{div}	Referans bölüm sayısı
n_{gen}	Maksimum yineleme sayısı
p	Değişken sayısı
P_c	Çaprazlama olasılığı
P_m	Mutasyon olasılığı
ϕ	GLM’de ölçek parametresi
$log(.)$	Logaritma fonksiyonu
$\mathbb{N}$	Doğal sayılar kümesi
$\mathbb{Q}$	Rasyonel sayılar kümesi
$\mathbb{R}$	Reel sayılar kümesi
θ	GLM’de konum parametresi
ub	Değişkenin aldığı üst sınır
$Var(.)$	Varyans
$\mathbb{Z}$	Tam sayılar kümesi

Kısaltmalar

EA	Evolutionary Algorithm (Evrimsel Algoritma)
IP	Integer Programming (Tam Sayılı Programlama)
IRLS	Iteratively Reweighted Least Squares (İteratifli Yeniden Ağırlıklandırılmış En Küçük Kareler)
GA	Genetic Algorithm (Genetik Algoritma)
GLM	Generalized Linear Models (Genelleştirilmiş Lineer Modeller)
LASSO	Least Absolute Shrinkage and Selection Operator (En Az Mutlak Büzülme ve Seçim Operatörü)
LP	Linear Programming (Doğrusal Programlama)
MAE	Mean Absolute Error (Ortalama Mutlak Hata)
MDNLP	Mixed Discrete Nonlinear Programming (Karma Kesikli Doğrusal Olmayan Programlama)
MDNSGA-II	Mixed Data Nondominated Sorting Genetic Algorithm-II (Karma Verili Baskın Sıralı Genetik Algoritma-II)
MDNSGA-III	Mixed Data Nondominated Sorting Genetic Algorithm-III (Karma Verili Baskın Sıralı Genetik Algoritma-III)
MILP	Mixed Integer Linear Programming (Karma Tam Sayılı Doğrusal Programlama)
MINLP	Mixed Integer Nonlinear Programming (Karma Tam Sayılı Doğrusal Olmayan Programlama)
MOEA	Multi Objective Evolutionary Algorithm (Çok Amaçlı Evrimsel Algoritma)
MOGA	Multi Objective Genetic Algorithm (Çok Amaçlı Genetik Algoritma)
MOO	Multi Objective Optimization (Çok Amaçlı Optimizasyon)
MSE	Mean Squared Error (Hata Kareler Ortalaması)
NLP	Nonlinear Programming (Doğrusal Olmayan Programlama)
NPGA	Niched Pareto Genetic Algorithm (Niş Pareto Genetik Algoritma)
NSGA	Nondominated Sorting Genetic Algorithm (Baskın Sıralı Genetik Algoritma)
NSGA-II	Nondominated Sorting Genetic Algorithm-II (Baskın Sıralı Genetik Algoritma-II)

NSGA-III	Nondominated Sorting Genetic Algorithm-III (Baskın Sıralı Genetik Algoritma-III)
OLS	Ordinary Least Squares (En Küçük Kareler)
RMSE	Root Mean Square Error (Ortalama Hata Kare Kökü)
SBX	Simulated Binary Crossover (Simüle Edilmiş İkili Çaprazlama)
SPEA	Strength Pareto Evolutionary Algorithm (Güçlü Pareto Evrimsel Algoritma)
SUR	Seemingly Unrelated Regression (Görünüşte İlişkisiz Regresyon)
VEGA	Vector Evaluated Genetic Algorithm (Vektör Hesaplamalı Genetik Algoritma)
VIF	Variance Inflation Factor (Varyans Artış Faktörü)

ŞEKİLLER DİZİNİ

Şekil 2.1 β 'nın (a) Yansız ve (b) Yanlı tahmin edicilerinin örnekleme dağılımı	20
Şekil 2.2 Varyans/Yanlılık ödünleşimi	24
Şekil 2.3 (a) Ridge, (b) LASSO ve (c) Elastik Net regresyon modelleri için tahmin ediciler	25
Şekil 3.1 Optimizasyon modellerinin sınıflandırılması	27
Şekil 3.2 Karar uzayı ve karar uzayına karşılık gelen amaç uzayı arasındaki ilişki	29
Şekil 3.3 (a) Başlangıç topluluğu ve Pareto yüzeyi, b) Pareto çözüm kümesi.....	31
Şekil 3.4 İki amaç fonksiyonu için olabilecek dört farklı durumda elde edilen Pareto çözüm kümeleri	31
Şekil 3.5 MOO yöntemlerinin sınıflandırılması.....	32
Şekil 3.6 Optimum arama ağacı örneği.....	38
Şekil 4.1 Kodlama ve kod çözme	45
Şekil 4.2 İkili kodlamada kullanılan çaprazlama yöntemleri.....	47
Şekil 4.3 GA operatörlerinde kullanılan başlıca yöntemler	49
Şekil 4.4 GA'nın akış şeması	56
Şekil 4.5 Pareto yüzeyi, basılgın ve baskın çözümler kümesi	58
Şekil 4.6 Baskın sıralama ve kalabalık uzaklık yaklaşımları	60
Şekil 4.7 NSGA-II'nin hızlı baskın sıralama ve kalabalık uzaklık hesaplama işleyişi	60
Şekil 4.8 NSGA-II'nin akış şeması	63
Şekil 4.9 $M = 3$ ve $p = 4$ iken bir problem için normalleştirilmiş bir referans düzleminde gösterilen 15 yapılandırılmış referans noktası	68
Şekil 4.10 Referans düzlemi	69
Şekil 4.11 Üç amaçlı problem için kesişim noktalarının hesaplanması ve ardından uç noktalardan hiperdüzlem oluşturma işlemi	70
Şekil 4.12 NSGA-III'ün normalleştirme işlemleri	71
Şekil 4.13 Popülasyon üyelerinin referans noktalarıyla ilişkisi	72
Şekil 4.14 F_l yüzeyinden $K = N - P_{t+1} $ bireyin seçimi.....	74
Şekil 4.15 NSGA-III'ün akış şeması	78
Şekil 4.16 (a) NSGA-II ve (b) NSGA-III yöntemlerinin çözümler arasındaki çeşitliliği korumak için kullandığı operatörler	80
Şekil 5.1 Uyarlanmış GA'nın akış şeması	86
Şekil 5.2 MDNSGA-II'nin akış şeması	90

Şekil 5.3 MDNSGA-III'ün akış şeması	93
Şekil 6.1 (a) Y_1 ve (b) Y_2 yanıt değişkenleri için kutu grafikleri	98
Şekil 6.2 (a) Y_1 ve (b) Y_2 yanıt değişkenleri için Q-Q grafikleri	98
Şekil 6.3 Y_1 yanıtıyla girdi değişkenleri arasındaki ilişkiyi gösteren saçılım grafikleri.....	99
Şekil 6.4 Y_2 yanıtıyla girdi değişkenleri arasındaki ilişkiyi gösteren saçılım grafikleri	100
Şekil 6.5 Girdi ve yanıt değişkenleri arasındaki korelasyon matrisi ısı haritası (Spearman sıra korelasyonu).....	101
Şekil 6.6 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için elde edilen Pareto çözüm kümeleri	105
Şekil 6.7. MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için elde edilen Pareto çözüm kümelerinin karşılaştırılması	106
Şekil 6.8 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümeleri	107
Şekil 6.9 MDNSGA-II ve MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümelerinin karşılaştırılması	108
Şekil 6.10 Yanıt değişkenleri arasındaki korelasyon matrisi ısı haritası	111
Şekil 6.11 (a) f_1 ve (b) f_2 amaç fonksiyonları için yüzey grafikleri.....	112
Şekil 6.12 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümeleri.....	114
Şekil 6.13 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması	114
Şekil 6.14 (a) f_1 ve (b) f_2 amaç fonksiyonları için yüzey grafikleri.....	115
Şekil 6.15 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için elde edilen Pareto çözüm kümeleri	116
Şekil 6.16 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması	116
Şekil 6.17 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümeleri.....	117
Şekil 6.18 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması	118

ÇİZELGELER DİZİNİ

Çizelge 2.1 Çok yanıtlı karma veri kümesi	12
Çizelge 2.2 GLM’de yaygın kullanılan bağ türleri, bağ fonksiyonu ve bağ fonksiyonunun tersi	14
Çizelge 2.3 GLM’de yaygın kullanılan üstel aile dağılımları, bağ ve varyans fonksiyonları	15
Çizelge 2.4 Farklı durumlarda uygulanan modelleme yöntemleri	26
Çizelge 4.1 Biyoloji ve genetik biliminde kullanılan terimlerin GA terminolojisindeki karşılıkları	43
Çizelge 4.2 GA için literatürde önerilen ayarlanabilir parametre değerleri	57
Çizelge 6.1 Enerji verimliliği karma veri kümesi	95
Çizelge 6.2 Enerji verimliliği karma veri kümesinin girdi ve yanıt değişkenleri, değişken türleri ve tanım kümeleri	96
Çizelge 6.3 Y_1 ve Y_2 yanıt değişkenleri için betimsel istatistikler	97
Çizelge 6.4 $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modelleri için elde edilen performans değerleri	103
Çizelge 6.5 MDNSGA-II’nin ayarlanabilir parametre değerleri	104
Çizelge 6.6 MDNSGA-III’ün ayarlanabilir parametre değerleri	104
Çizelge 6.7 Enerji verimliliği karma verisinde $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için MDNSGA-II ile elde edilen Pareto çözüm kümesi	105
Çizelge 6.8 Girdi değişkenlerinin gerçek ve kodlanmış düzeyleri	109
Çizelge 6.9 Deneysel çalışma sonucu elde edilen kodlanmış girdi değişken düzeyleri ve gözlenen yanıt değişken değerleri	109
Çizelge 6.10 Gıda karma verisi için girdi değişkenleri, değişken türü ve tanım kümesi	110
Çizelge 6.11 Gıda karma verisi için yanıt değişkenlerinin Shapiro-Wilk Normallik testi	110
Çizelge 6.12 MDNSGA-II için ayarlanabilir parametreler ve değerleri	113
Çizelge 6.13 MDNSGA-III için ayarlanabilir parametreler ve değerleri	113
Çizelge 6.14 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen optimal girdi değişken değerleri ve amaç fonksiyonu değerleri	115

1. GİRİŞ VE ÖNCEKİ ÇALIŞMALAR

1.1 Giriş

Fen, sağlık, sosyal, tıp, mühendislik, finans vb. alanlarda yapılan araştırmalarda kullanılan veriler, hem nicel hem de nitel değişken türlerini birlikte içerebilir. Veri kümesinin nicel ve nitel değişkenleri birlikte içermesi durumunda *karma veri (mixed data)* ortaya çıkar (Chavent vd. 2017). Nitel değişkenler kategorik değişkenler olarak da bilinir. Kategorik değişkenler sınıflandırılarak kesikli nicel değerlerle ifade edildiğinde ise ilgilenilen problem, kesikli ve sürekli nicel değer alan değişkenlere sahip *karma veri problemleri* olarak adlandırılır. Verinin işlenebilmesi ve analiz edilebilmesi için veri yapısının araştırmacı tarafından iyi anlaşılmış olması gerekir. Gözlem ve deney yoluyla elde edilen veri kümesinde birden fazla girdi (bağımsız, açıklayıcı) ve yanıt (bağımlı, açıklanan) değişken(ler)i yer alabilir. Problem, girdi değişkenleri bakımından karma olabileceği gibi yanıt değişkenleri bakımından da karma yapıda olabilir.

Disiplinlerarası çalışmalarda, karma veriyle ilgili istatistiksel modelleme ve optimizasyon çalışmaları oldukça önemli yer tutar. Oluşturulacak istatistiksel modelin güvenilirliği optimizasyon sonuçlarının güvenilirliğini doğrudan etkiler. Bu nedenle, oluşturulacak modelin veri kümesini iyi temsil etmesi gerekir. Karma veri kümesindeki yanıt değişken(ler)inin sürekli ya da kesikli değer almasına göre de oluşturulacak model(ler) farklılık gösterir. Yanıt değişkeni, sürekli değerli iken çoklu doğrusal regresyon analizi, kesikli değerli iken dağılım yapısına uygun olarak lojistik, binom veya Poisson regresyon analizi uygulanabilir. Değişkenler arasındaki doğrusal ilişkiyi modellemek için kullanılan doğrusal regresyon modelleri, Genelleştirilmiş Doğrusal Modellerin (Generalized Linear Models -GLM) özel halleridir. GLM, kitle ortalamasının bir bağ (link) fonksiyonuyla doğrusal tahmin ediciye bağlı olmasına izin veren doğrusal modellerin genelleştirilmiş halidir (Garrido ve Zhou 2009). GLM kullanılarak tahmini yanıt fonksiyonu elde edilir. Tahmini yanıt fonksiyonun elde edilmesiyle modelleme aşaması tamamlanır.

GLM ile elde edilen tahmini yanıt fonksiyonu amaç fonksiyonu olarak ele alınarak karma veri modelinin optimizasyonu aşamasına geçilir. Optimizasyon en genel anlamda, bir

amaç fonksiyonu için belirli kısıtlar altında optimal çözümün elde edilmesidir (Apaydın 2002). Veri kümesinde yanıt değişkeninin birden fazla olması durumunda problem, çok amaçlı optimizasyon (Multi Objective Optimization-MOO) problemi olarak adlandırılır. MOO problemlerinde her bir yanıt fonksiyonu bir amaç fonksiyonu olarak ele alınarak eş anlı olarak optimize edilir. Amaç sayısı birden fazla ve amaçların birbiri ile çelişmesi durumunda birbirine ne baskın ne de basılın çözümler olarak adlandırılan *Pareto çözüm kümesi* ile MOO problemlerinin çözümleri elde edilir. Karar verici, elde edilen Pareto çözüm kümesinden, ilgilendiği tüm amaçları tatmin edecek uzlaşık çözümü seçebilir.

Karma veri içeren optimizasyon problemleri, kesikli değerler alan karar değişkenleri içerdiği için kombinatoriyaldır. Bilgisayar alanındaki hızlı gelişmelere rağmen, bu tür problemlerin çözümleri uzun zaman ve çok maliyet gerektirir. Bu nedenle büyük boyutlu ve karmaşık problemlerin çözümünde makul zamanda optimal sonuca yakın çözümler üreten sezgisel (heuristic) yöntemlerin kullanımı ön plana çıkmaktadır. Sezgisel optimizasyon yöntemlerinden biri olan Genetik Algoritma (Genetic Algorithm-GA), karma optimizasyon problemini çözmek için etkili ve sağlam bir yöntemdir (Wu ve Chow 1994). Sezgisel yöntemler literatürde yapay zeka yöntemleri olarak da adlandırılmaktadır (Karaboğa 2020). Yapay zeka optimizasyon algoritması olarak tanımlanan GA, esnek hesaplama yöntemleri arasında önemli bir yer tutar.

MOO problemlerini çözmek için geliştirilmiş evrimsel algoritmalar (Evolutionary Algorithms-EA) sınıfında yer alan çok amaçlı GA (Multi Objective Genetic Algorithm-MOGA) ile birden fazla amaç fonksiyonuna sahip optimizasyon problemleri için Pareto çözüm kümesi bulunur. Pareto tabanlı yöntemler sınıfında yer alan MOGA, en çok baskın sıralı GA yöntemine dayanan Baskın Sıralı GA-II (Nondominated Sorting Genetic Algorithm-II / NSGA-II) yöntemini kullanır (Cebeci 2021). NSGA-II'nin diğer MOGA'lara göre üstünlüğü, hızlı baskın sıralama (fast nondominated sorting) ve kalabalık uzaklığı (crowding distance) yaklaşımlarıdır. Bu yaklaşımlarla, baskın yüzey oluşturulup Pareto çözüm kümesinde alternatif çözümler elde edilerek Pareto yüzeye daha kısa sürede ulaşılır (Türkşen 2011). NSGA-II ile Pareto yüzeyde yeterince çeşitlilik sağlanamadığında, referans noktalarına dayalı yeni çeşitlilik oluşturma yöntemi olan NSGA-III kullanılabilir. Pareto tabanlı algoritmalarından biri olan NSGA-III'te, NSGA-

II'deki kalabalık uzaklık operatörü bir kümeleme operatörüyle değiştirilip çözüm çeşitliliğini garanti etmek için iyi dağılmış referans noktaları kümesi kullanılır (Deb ve Jain 2014).

Bu tez çalışmasında, sürekli, kesikli ve tam sayılı değerler alan girdi değişkenlerine sahip çok yanıtli karma veri kümelerinin istatistiksel olarak modellenmesi ve bu modellerin bir amaç fonksiyonu biçiminde ele alınarak çok amaçlı optimizasyonunun esnek hesaplama yöntemleriyle yapılması amaçlanmıştır. Karma veri kümesindeki yanıtların tahmini modellerinin elde edilmesinde literatürde mevcut olan GLM, Ridge, LASSO, Elastik Net ve SUR yöntemleri uygulanmıştır. MOO aşamasında esnek hesaplama yöntemleri olarak literatürde sıklıkla kullanılan GA tabanlı NSGA-II ve NSGA-III tercih edilmiştir. Tez çalışmasının temel hedefi olan karma verili problemlerin çok amaçlı optimizasyonu için bu esnek hesaplama yöntemlerinin karar değişkenlerinin gösteriminde, başlangıç popülasyonunun oluşturulmasında ve genetik operatörlerin uygulanması adımlarında uyarlamalar yapılmıştır. Uyarlanan yöntemler tez çalışması kapsamında, MDNSGA-II ve MDNSGA-III olarak adlandırılmıştır.

Çalışmanın ikinci bölümünde, GLM hakkında bilgi verilip karma verilerin modellenmesi aşamasında girdi değişkenleri ilişkiliyken kullanılan Ridge, LASSO ve Elastik Net regresyonu ve yanıt değişkenleri ilişkiliyken kullanılan Görünüşte İlişkisiz Regresyon (Seemingly Unrelated Regression-SUR) anlatılacaktır.

Çalışmanın üçüncü bölümünde, optimizasyon modellerinin tanımı verilerek özelliklerine ve çözüm yollarına kısaca değinilecektir. Amaç fonksiyonu sayısına, fonksiyon yapılarına ve değişken türlerine göre optimizasyon modelleri tanıtılacaktır. Karma veri içeren optimizasyon problemlerinin çözümü için kullanılan yöntemlerden söz edilecektir.

Çalışmanın dördüncü bölümünde, esnek hesaplama yöntemleri arasında önemli bir yer tutan GA detaylı olarak anlatılacaktır. Tek amaçlı GA'nın genel işleyişinden bahsedilip çok amaçlı optimizasyonda uygulanan NSGA-II ve NSGA-III yöntemleri tanıtılacaktır.

Çalışmanın beşinci bölümünde, sürekli, kesikli ve tam sayılı değer alan değişkenlere sahip karma veri modellerinin tek amaçlı GA ile optimizasyonu anlatılacaktır. Karma veri modellerinin çok amaçlı optimizasyonunda ise yapay zeka optimizasyon algoritmaları olan NSGA-II ve NSGA-III yöntemlerinde yapılacak uyarlamalarla Pareto çözüm kümesinin elde edilmesi amaçlanmıştır. GA'nın değişken gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin uygulanması aşamalarında yapılacak olan uyarlamalar ayrıntılı olarak verilecektir. Karma verili problemlerin optimizasyonu için gerekli hesaplama süresini kısaltacak değer kodlama kullanılarak NSGA-II ve NSGA-III yöntemleri uyarlanıp karma verili NSGA-II (MDNSGA-II) ve karma verili NSGA-III (MDNSGA-III) yöntemleri oluşturulmuştur. MDNSGA-II ve MDNSGA-III ile kesikli değişkenlerin tanım kümesinden kesikli değer alması sağlatılarak çok amaçlı karma verili problemler için Pareto çözüm kümesi elde edilecektir. Tez çalışmasının özgünlüğünü oluşturan bu bölümde, MDNSGA-II ve MDNSGA-III esnek hesaplama yöntemlerinin işleyişi detaylı biçimde anlatılarak akış şemalarına yer verilecektir.

Çalışmanın altıncı bölümünde, uygulama amacıyla UCI Machine Learning Repository veri tabanından enerji verimliliği konulu karma veri kümesine ve gıda alanında literatürde mevcut olan deneysel veri kümesi karma verili biçimde ele alınarak oluşturulan gıda karma veri kümesine bu tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleri uygulanacaktır. Uygulanan yöntemlerle karma verili çok amaçlı problemler için Pareto çözüm kümeleri elde edilerek Pareto çözüm kümelerindeki çözümler değerlendirilecektir.

Çalışmanın yedinci bölümünde, analizlerden elde edilen sonuçların değerlendirilmesi yapılacaktır.

1.2 Önceki Çalışmalar

Bu bölümde, karma veri modellerinin optimizasyonu için yapılan önceki çalışmalardan söz edilecektir.

Rajeev ve Krishnamoorthy (1992), yapısal sistemleri kesikli tasarım değişkenleriyle optimize etmek için kısıtlı probleme ceza tabanlı bir dönüşüm uygulayarak kısıtsız probleme dönüştürüp GA'yı kullanmışlardır.

Lin ve Hajela (1992), karma kesikli doğrusal olmayan programlama (Mixed Discrete Nonlinear Programming-MDNLP) problemlerine GA yöntemini uygulamışlardır. GA'da değişken kodlaması yapıldığından karma değişkenler içeren problemlere GA'nın uygulanmasının ideal olduğunu belirtmişlerdir. GA'yı Dal-Sınır yöntemine alternatif olarak önermişlerdir. GA'nın geleneksel gradyan tabanlı yöntemlere kıyasla global optimumun yerini bulma konusunda daha iyi olduğunu belirtmişlerdir.

Arora ve Huang (1994), Huang ve Arora (1995) araştırma makalelerinde MDNLP problemlerini çözmek için kullanılan yöntemleri altı kategoride sınıflandırmışlardır: Dal-Sınır Yöntemi, Benzetimli Tavlama, Sıralı Doğrusal Programlama, Ceza Fonksiyonu Yaklaşımı, Lagrange Gevşemesi ve Yuvarlama, Kesme Düzlemi, Sezgisel Yöntemler, GA vb.

Wu ve Chow (1994), mühendislik tasarımında ortaya çıkan MDNLP problemlerine sezgisel kombinatoriyal optimizasyon tekniği olan GA'yı uygulayarak GA'nın bu tür problemleri çözmede etkili ve sağlam bir yöntem olduğunu göstermişlerdir.

Wu ve Chow (1995a), kafes kirişlerinin kesikli optimizasyonu için GA'yı kullanmışlardır. Kesikli değişkenleri, işaretli ondalıklı tam sayılar aracılığıyla ikili rakam dizileri $\{0,1\}$ ile kesikli değerler arasında eşleme yaparak ele almışlardır. Uygulama sonuçlarında GA'nın verimliliğini ve sağlamlığını göstermişlerdir.

Lin vd. (1995), MDNLP problemlerini çözmek için bilinen GA'da önemli değişiklikler yapmışlardır. Yapılan bu değişiklikler; değişken gösterimi, genetik parametrelerin ayarlanması, başlangıç popülasyonunun oluşturulması, genetik operatörlerin uygulanması ve uygunluk fonksiyonlarının hesaplanması aşamalarında. Değişken gösteriminde kromozom uzunluğunu önemli ölçüde azalttığı için hibrit dizi gösterimini kullanmışlardır. Hibrit dizi gösteriminde, her bit konumu için 0 ve 1 yerine 0'dan 9'a kadar olan tam sayıları kullanmışlardır.

Li ve Gen (1996), Ghasemi vd. (1997), karma tam sayılı doğrusal olmayan programlama (Mixed Integer Nonlinear Programming-MINLP) problemlerini çözmek için ceza fonksiyonu ve ikili kodlama kullanarak GA uygulamışlardır.

Xiong ve Rao (2004), bulanık formülasyonu hibrit GA ile birleştiren sentetik bir karma kesikli bulanık doğrusal olmayan lineer programlama problemi için optimizasyon yöntemi sunmuşlardır. Bu yaklaşımla, tatmin edici ve güvenilir sonuçların elde edilebileceğini göstermişlerdir.

Rao ve Xiong (2005a), MDNLP problemlerinin çözümü için yeni bir hibrit GA sunmuşlardır. Küresel optimum noktayı içeren optimum uygulanabilir bölgeyi GA ile belirleyip, kesikli tek boyutlu aramayla birleştirilmiş hibrit negatif alt gradyan yöntemini de son optimum çözümü bulmak için kullanmışlardır.

Deep vd. (2009), MINLP problemlerini çözmek için Laplace çaprazlama (Laplace Crossover-LX) ve güç mutasyonunu (Power Mutation-PM) karar değişkenlerinin bir kısmı veya tamamı üzerindeki tam sayı kısıtlamalarını ele alacak şekilde değiştirerek karma tam sayılı Laplace çaprazlama güç mutasyonu olarak adlandırdıkları algorithmada değer kodlu GA uygulamışlardır. Bu algorithmada, Laplace çaprazlamasını ve Güç mutasyonunu tam sayılı değişkenler için değiştirerek kullanmışlardır. Önerdikleri algoritma ile karar değişkenleri üzerindeki tam sayı kısıtlamasının karşılanması için özel bir kesme prosedürü ve kısıtlamanın ele alınması için parametreden bağımsız bir ceza yaklaşımı kullanmışlardır.

Song vd. (2021) hem sürekli hem de kesikli optimizasyon problemlerinin çözümünde farklı operatörlerin ve parametre ayarlarının etkisini araştırmak amacıyla GA üzerinde kapsamlı deneyler gerçekleştirmiştir. Deneylerin sonucunda, turnuva seçimi, tekdüze çaprazlama ve yeni bir kombinasyona dayalı mutasyonun sürekli problemler için en iyi seçim olduğunu gösterirken, rulet çarkı seçimi, mesafeyi koruyan çaprazlama ve takas mutasyonun kesikli problemler için en iyi seçim olduğunu belirtmiştir.

Ebrahimi vd. (2021), en düşük spesifik enerji tüketimini elde etmek için MR kriyojenik sürecin optimizasyonunu araştırmak için GA'yı kullanmışlardır. Çok değişkenli optimizasyon probleminin matematiksel zorluklarını çözmek için beş adımdan oluşan kesikli sürekli bir GA geliştirmişlerdir. Önerdikleri optimizasyon algoritmasını kesikli arama ve sürekli arama olmak üzere iki bileşene ayırmışlardır.

Literatürde çok amaçlı MDNLP problemlerinin çözümü için yapılan çalışmalar aşağıdaki gibi özetlenebilir.

Hajela ve Shih (1989), çok amaçlı MDNLP problemlerine çözümler elde etmek için değiştirilmiş global kriter yaklaşımının bir min-max varyantını önermişlerdir. Bu yaklaşımı, karma veriler için yapılandırılan Dal-Sınır Yöntemi ile birlikte kullanmışlardır.

Wang vd. (2005), bir bina tasarımında, tasarımcılara yardımcı olabilecek MOO problemini hem ekonomik hem de çevresel kriterler bakımından değerlendirerek Pareto optimal çözümün belirlenmesi için MOGA kullanmışlardır.

Rao ve Xiong (2005b), çok amaçlı karma kesikli bulanık doğrusal olmayan programlama problemlerini çözmek için bulanık λ formülasyonu ve oyun teorisi tekniklerinin karma kesikli hibrit genetik algoritma ile birleştirildiği yeni bir yöntem sunmuşlardır. Önerdikleri bu hibrit yöntemle, belirsiz bir ortamda tatmin edici ve gerçekçi sonuçlar elde edilebileceğini belirtmişlerdir.

Ahmadi vd. (2013), havza ölçeğinde tarımsal koruma uygulamalarının hedeflenen şekilde uygulanması için ikili ve karma değişkenli NSGA-II yöntemlerinin hibridizasyonuna yönelik bir yöntem geliştirmişlerdir.

El-Kribi vd. (2013), sürekli ve kesikli değişkenlere sahip dört çubuklu bir sistemin mekatronik tasarımını eş anlı optimize etmek için NSGA-II'yi kullanarak farklı tasarım koşulları için analiz etmişlerdir.

Tong vd. (2014), karma kesikli problemler için daha önce geliştirilen ve erken parçacık kümelenmesini önlemek için özel bir çeşitlilik koruma tekniği içeren karma kesikli parçacık sürü optimizasyonu algoritmasını MOO problemlerini çözmek için kullanmışlardır.

Kondoh vd. (2016), değer kodlu GA'lar üzerinde kesikli değişkenlerin, Pareto optimal çözümlerin yakınsamasını ve çeşitliliğini nasıl etkileyebileceğini açıklamışlardır. Her değişken için değişken değerinin virgülden sonraki anlamlı basamak sayısını değiştirip, NSGA-II kullanmışlardır. Gerçek mühendislik problemlerini gerçekçi bir şekilde temsil etmek için değer kodlama kullanarak değişken değerinin virgülden sonraki anlamlı basamak sayısını ayarlamanın kritik öneme sahip olduğunu belirtmişlerdir.

Holzman ve Smith (2018), kesikli MOO problemine tam verimli bir çözüm kümesi üretmek için modifiye edilmiş arttırılmış ağırlıklı Tchebychev normunu sunmuşlardır. Kullandıkları algoritmanın çalışma sürelerini literatürde önerilen algoritmaların çalışma süreleriyle karşılaştırmışlardır.

Guangyong vd. (2018), belirsizlikler içeren mühendislik yapılarının tasarımı için yeni çok amaçlı kesikli robust optimizasyon algoritması önermişlerdir. Çalışmalarında çok kriterli karar verme tekniğini Taguchi yöntemine dahil ederek çok amaçlı kesikli sağlam tasarımı ele almayı amaçlamışlardır. Önerdikleri algoritmanın NSGA-II ile elde edilen Pareto sınırlarına yakın olduğunu belirtmişlerdir.

Roy vd. (2021), iki dallı bir genetik algoritmayı küresel bir arama aracı olarak yerel arama için gradyan tabanlı bir yaklaşımla birleştiren, kısıtlı çok amaçlı karma kesikli doğrusal olmayan programlama problemine çözüm bulmak için yeni bir hibrit yaklaşım önermişlerdir.

Jaber vd. (2022), çok amaçlı MINLP problemlerinin yaklaşık Pareto yüzeyini geliştirmek için çok kriterli Dal-Sınır Yöntemi ile NSGA-II'ye dayanan genel bir hibrit yaklaşım önermişlerdir. Önerdikleri Dal-Sınır NSGA-II algoritmasının çok amaçlı karma tam sayılı doğrusal olmayan programlama problemlerini çözmek için rekabetçi bir alternatif olabileceğini göstermişlerdir.

Tavana vd. (2016), NSGA-III ve çok amaçlı Parçacık Sürü Optimizasyonu yöntemlerini kullanarak X-çubuğu kontrol şeması tasarım problemini optimize etmeye çalışmışlardır. Her iki algoritmada, tam sayılı ve sürekli karar değişkenleriyle üç hedefi ve iki kısıtlamayı dikkate alacak şekilde uyarlamalar yaparak elde ettikleri Pareto optimal sınırları karşılaştırmak için performans ölçütleri kullanmışlardır. NSGA-III'ün performans ölçütlerine göre daha iyi sonuçlar elde ettiğini ve NSGA-III algoritmasının çok amaçlı Parçacık Sürü Optimizasyonu algoritmasına göre daha fazla baskın çözüm ürettiğini ileri sürmüşlerdir.

Literatürde MOO problemlerinin çözümü için NSGA-II ve NSGA-III kullanılarak yapılan çalışmalar aşağıdaki gibi özetlenebilir.

Ciro vd. (2016), üç amacın aynı anda en aza indirilmesinin amaçlandığı çok amaçlı bir problemi ele alarak küçük boyutlu örnekler için NSGA-II ve NSGA-III'ün benzer bir performansa sahip olduğunu, ancak NSGA-III'ün NSGA-II'den biraz daha yüksek bir hiperhacim ortalaması sunduğunu belirtmişlerdir. Büyük boyutlu örnekler için ise problemin boyutu arttığında NSGA-III'ün, NSGA-II'den daha iyi bir performansa sahip olduğunu göstermişlerdir.

Ishibuchi vd. (2016), NSGA-III'ün performansını NSGA-II ile karşılaştırmak için en az üç ve en fazla on amaç fonksiyonuna sahip, dört farklı, çok amaçlı test problemleri üzerinde çalışmışlardır. Çalışmalarında NSGA-III'ün her zaman NSGA-II'den daha iyi performans göstermediğini ifade etmişlerdir. Çok amaçlı sırt çantası problemleri için NSGA-II'nin daha iyi sonuçlar verdiğini görmüşler ve test problemlerinin seçiminin performans karşılaştırma sonuçları üzerinde amaç fonksiyonu sayısından daha büyük bir etkiye sahip olduğunu belirtmişlerdir.

Kanata vd. (2019), gerçek güç kayıplarını ve gerilim sapmalarını en aza indirmeyi amaçladıkları doğrusal ve dışbükey olmayan, çok kısıtlı ve karma değişkenli özelliklere sahip çok amaçlı optimal reaktif güç dağıtım problemlerinde, NSGA-III yöntemini kullanmışlardır. Simülasyon sonuçlarını, karma değişkenleri uygulayan önceki çalışmalarla karşılaştırıp önerdikleri optimizasyon yönteminin, çok amaçlı geliştirilmiş Parçacık Sürü Optimizasyonu, çok amaçlı Parçacık Sürü Optimizasyonu ve çok amaçlı Karınca Aslanı Optimizasyonuna kıyasla, gerçek güç kayıplarını ve hesaplama süresini en aza indirmede daha verimli ve güvenilir olduğunu göstermişlerdir.

Tanabe ve Ishibuchi (2020), hedef sayısı, Pareto yüzeyinin şekli ve tasarım değişkenlerinin türü açısından çeşitli problemleri içeren 16 sınırlı-kısıtlı gerçek dünya probleminin oluştuğu MOO problemi paketi sunmuşlardır. Bu problem paketinde 4'ü çok amaçlı karma tam sayılı olmak üzere 16 optimizasyon probleminin optimizasyon sonucunu, NSGA-II, ayrıştırmaya dayalı çok amaçlı evrimsel algoritma (Multi Objective Evolutionary Algorithm-MOEA), ayrıştırmaya dayalı MOEA yönteminde kullanılan ceza sınırı kesişimi yöntemi, NSGA-III, çok amaçlı aramada gösterge tabanlı seçim ve baskın hiperhacime dayalı çok amaçlı seçim yöntemleriyle karşılaştırmışlardır.

Yannibelli vd. (2020), MOEA otomatik ölçekleyici adı verilen son teknoloji ürünü bir otomatik ölçekleyiciye odaklanan çalışma yapmışlardır. Yaptıkları çalışmada, NSGA-II ve NSGA-III'ü kullanarak bu optimizasyon algoritmalarının MOEA otomatik ölçekleyicinin performansı üzerindeki etkilerini analiz etmişlerdir. Gerçek dünya deneyleriyle yaptıkları simülasyon çalışmalarında NSGA-II yerine NSGA-III kullanıldığında MOEA'nın önemli ölçüde iyileştiğini belirtmişlerdir.

Chaudhari vd. (2022), bir Stiren reaktörünün üç amaçlı optimizasyon problemi için NSGA-III'ün etkinliğini NSGA-II ile karşılaştırmıştır. NSGA-III ile elde edilen sonuçların NSGA-II ile elde edilen sonuçlara göre daha çeşitli optimum çalışma koşullarını sağladığını göstermişlerdir.

Santos vd. (2024), NSGA-II ve NSGA-III algoritmalarının iki amaçlı atölye tipi çizelgeleme problemlerini çözmedeki performanslarını karşılaştırmıştır. Karşılaştırma amacıyla, hiperhacim performans ölçütünün sonuçlarını istatistiksel olarak analiz etmişlerdir. Elde ettikleri sonuçlar, bu iki algoritmanın performansının önemli ölçüde farklı olmadığını ve bu nedenle NSGA-III'ün iki amaçlı atölye tipi çizelgeleme problemlerini çözmede açık bir avantaj sağlamadığını göstermiştir.

2. İSTATİSTİKSEL MODELLEME

İstatistiksel modelleme ile yanıt ve girdi değişken(ler)i arasındaki ilişkiler belirlenebilmekte ve veri kümesindeki ham verilerden bilgi elde edilerek bu bilgiler doğrultusunda tahminler yapılabilmektedir. İstatistiksel modellemede kullanılan yöntemler, araştırmanın amacına ve kullanılan değişkenlerin türlerine göre farklılık gösterir. Güvenilir bir model kurmak için veri yapısının iyi anlaşılması gerekir. Çünkü, doğru analizler veri yapısının iyi anlaşılmasıyla elde edilir. Yanıt değişkenlerinin birden fazla olması durumunda çok yanıtlı verilerin analizi için yapılan modelleme çalışmalarında, girdi değişkenlerinin $(X_i, i = 1, 2, \dots, p)$ ilgilenilen yanıt değişkenleri $(Y_j, j = 1, 2, \dots, r)$ üzerindeki etkisi değişkenler arasındaki ilişki dikkate alınarak incelenmelidir. Çizelge 2.1’de, p tane girdi değişkeni, r tane yanıt değişkenine sahip n gözlemlili çok yanıtlı karma veri kümesi örneği yer almaktadır. Burada, girdi değişkenleri matrisi $X = [x_{ij}]_{n \times p}$ ve i .yanıt değişkenine ilişkin gözlenen değerler vektörü $[y_{1i} \ y_{2i} \ \dots \ y_{ni}]$, $i = 1, 2, \dots, r$ ’dir.

Çizelge 2.1 Çok yanıtlı karma veri kümesi

No	Girdi Değişkenleri				Yanıt Değişkenleri			
	X_1	X_2	...	X_p	Y_1	Y_2	...	Y_r
1	x_{11}	x_{12}	...	x_{1p}	y_{11}	y_{12}	...	y_{1r}
2	x_{21}	x_{22}	...	x_{2p}	y_{21}	y_{22}	...	y_{2r}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	x_{n1}	x_{n2}	...	x_{np}	y_{n1}	y_{n2}	...	y_{nr}

2.1 Genelleştirilmiş Doğrusal Modeller

Genelleştirilmiş Doğrusal Modeller (Generalized Linear Models-GLM), kesikli ve sürekli yanıt değişkenleri için birleştirilmiş regresyon modelleri sınıfıdır (Khuri vd. 2006). GLM, yanıt değişkeni dağılımının Normal dağılım varsayımını sağlamadığı

durumları da göz önüne alan hem doğrusal hem de doğrusal olmayan regresyon modellerinin bir birleşimidir (Collins 2008). Başka bir deyişle GLM, kitle ortalamasının bir bağ fonksiyonu ile doğrusal tahmin ediciye bağlı olmasına izin veren doğrusal modellerin genelleştirilmiş halidir (Garrido ve Zhou 2009). İlk olarak Nelder ve Wedderburn (1972) tarafından kullanılan GLM, üç temel bileşenden meydana gelmektedir. Bu bileşenler, doğrusal tahmin edicilerin bulunduğu sistematik kısım, yanıt değişkeninin dağılımı ve bağ fonksiyonudur. Modelin sistematik kısmı, p tane girdi değişkeni $(X_1, X_2, \dots, X_p)$ girdi değişkeni içerir. Yanıt değişkeni ile girdi değişkenleri arasındaki fonksiyonel ilişki GLM ile elde edilir. GLM’de yanıt değişkeninin dağılımı, Normal, Poisson, Binom, Üstel ve Gamma dağılımlarını içeren üstel ailenin bir üyesi olmalıdır. Yanıt değişkeninin sürekli olduğu durumlarda Gamma, Normal ve Ters Normal dağılım, kesikli olduğu durumlarda Poisson, Bernoulli ve Binom dağılımı üstel aile örnekleridir (Anderson vd. 2012). GLM’de, Y ’nin dağılımı üstel ailenin genel formunda

$$f_Y(y; \theta, \varphi) = \exp \left\{ \frac{y\theta - b(\theta)}{a(\varphi)} + c(y, \varphi) \right\} \quad (2.1)$$

biçiminde yazılır (McCullagh ve Nelder 1989). Burada, $a(\bullet)$, $b(\bullet)$ ve $c(\bullet)$ bazı özel fonksiyonlardır. θ , dağılım ortalamasının (μ) bir fonksiyonu olup dağılımın konumu hakkında bilgi verir ve *kanonik parametre* olarak adlandırılır. φ , dağılımın ölçeği hakkında bilgi verir ve yayılım parametresidir. $c(y, \varphi)$, gözlemler ile yayılım parametresinin bir fonksiyonudur. Üstel dağılım ailesinin olabilirlik fonksiyonunun logaritmasının θ ’ya göre birinci ve ikinci türevleri alınıp sıfıra eşitlendiğinde, Y ’nin ortalama ve varyansı sırasıyla

$$E(Y) = \mu = b'(\theta) \quad (2.2)$$

ve

$$Var(Y) = b''(\theta)a(\varphi) \quad (2.3)$$

olur. GLM’de temel amaç, yanıt değişkeninin beklenen değerinin uygun bir fonksiyonu için bir model geliştirmektir.

Bağ fonksiyonu $g(\bullet)$, yanıt değişkeninin ortalaması ile doğrusal tahmin ediciler arasında ilişki kurulmasını sağlar. η olarak da gösterilen bu bağ fonksiyon

$$g(\mu) = \eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = \mathbf{X}^T \boldsymbol{\beta} \quad (2.4)$$

biçiminde ifade edilir. Burada, $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \dots \ \beta_p]^T, (p+1) \times 1$ boyutlu regresyon katsayı vektörüdür (Öztürk 2012). Eşitlik (2.4) ile ifade edilen bağ fonksiyonu $g(\bullet)$, monoton ve türevlenebilir olmalıdır. Bu fonksiyonun birebir olduğu anlamına gelir. Bu nedenle bağ fonksiyonunun ters dönüşümü vardır. Ortalama fonksiyon da denilen $g^{-1}(\bullet)$ ters fonksiyonu $\mu = g^{-1}(\eta) = g^{-1}(\mathbf{X}^T \boldsymbol{\beta})$ 'dır (Friendly ve Meyer 2015). Çizelge 2.2'de, GLM'de yaygın kullanılan bağ fonksiyonu ve bağ fonksiyonunun tersi verilmiştir.

Çizelge 2.2 GLM'de yaygın kullanılan bağ türleri, bağ fonksiyonu ve bağ fonksiyonunun tersi (Fox 2015)

<i>Bağ Türü</i>	<i>Bağ Fonksiyonu</i>	<i>Bağ Fonksiyonunun Tersi</i>
<i>Özdeş</i>	μ_i	η_i
<i>Kare-kök</i>	$\sqrt{\mu_i}$	η_i^2
<i>Log</i>	$\log_e(\mu_i)$	$\exp(\eta_i)$
<i>Ters</i>	μ_i^{-1}	η_i^{-1}
<i>Ters-kare</i>	μ_i^{-2}	$\eta_i^{-1/2}$
<i>Logit</i>	$\log_e \frac{\mu_i}{1 - \mu_i}$	$\frac{1}{1 + \exp(-\eta_i)}$
<i>Probit</i>	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$
<i>log-log</i>	$-\log_e[-\log_e(\mu_i)]$	$\exp[-\exp(-\eta_i)]$
<i>Tamamlayıcı log-log</i>	$\log_e[-\log_e(1 - \mu_i)]$	$1 - \exp[-\exp(\eta_i)]$

* μ_i , yanıt değişkeninin beklenen değeri, η_i , doğrusal tahmin edici ve $\Phi(\bullet)$ standart normal dağılımın kümülatif dağılım fonksiyonudur.

Bağ fonksiyonunun özdeş olması durumunda dağılım Normal olur ve değişkenler arasındaki ilişki klasik bilinen regresyon analizi ile oluşturulur. Çoklu doğrusal regresyon modelinin en genel hali

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (2.5)$$

biçimindedir. Burada, β_j , $j = 0, 1, \dots, p$ parametreleri regresyon katsayıları ve ε hata terimidir. Değişkenler arasındaki ilişkiyi bir polinomsal denklem aracılığıyla ifade etmede kullanılan polinomsal regresyon, çoklu doğrusal regresyonun özel bir durumudur. Polinomsal modeller, etkili ve esnek bir eğri uydurma tekniğidir (Ostertagová 2012). p değişkenli ikinci dereceden bir polinomsal model

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \beta_{ii} X_i^2 + \sum_{i=1}^p \sum_{i < j}^p \beta_{ij} X_i X_j + \varepsilon \quad (2.6)$$

biçiminde verilen doğrusal regresyon modelidir. Eşitlik (2.6)'daki model parametreleri En Küçük Kareler (Ordinary Least Squares-OLS) yöntemi kullanılarak tahmin edilebilir. Parametreler tahmin edildiğinde elde edilen tahmini yanıt modeli

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^p \hat{\beta}_i X_i + \sum_{i=1}^p \hat{\beta}_{ii} X_i^2 + \sum_{i=1}^p \sum_{i < j}^p \hat{\beta}_{ij} X_i X_j \quad (2.7)$$

olur. Çizelge 2.3'te GLM'de yaygın kullanılan üstel aile dağılımları ve bu dağılımlarda kullanılan bağ fonksiyonları verilmiştir.

Çizelge 2.3 GLM'de yaygın kullanılan üstel aile dağılımları, bağ ve varyans fonksiyonları (Faraway 2006)

<i>Dağılım</i>	<i>Bağ fonksiyonu</i>	<i>Varyans fonksiyonu</i>
<i>Binom</i>	$\eta = \log\left(\frac{\mu}{1-\mu}\right)$ (logit bağ)	$\mu(1-\mu)$
<i>Poisson</i>	$\eta = \log(\mu)$ (log bağ)	μ
<i>Normal</i>	$\eta = \mu$ (özdeş bağ)	1
<i>Gamma</i>	$\eta = \mu^{-1}$ (ters bağ)	μ^2
<i>Ters Normal</i>	$\eta = \mu^{-2}$ (ters kare bağ)	μ^3

GLM'de kullanılan Binom, Poisson, Normal ve Gamma dağılımları için oluşturulan bağ fonksiyonları aşağıda verilmiştir.

Binom Dağılımı: Y yanıt değişkeni Binom dağılımlı olsun, $Y \sim B(n, p), n \in \mathbb{Z}^+, p \in (0,1)$. Y 'nin olasılık fonksiyonu (of)

$$f_Y(y; p) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y = 0, 1, 2, \dots, n \quad (2.8)$$

olup eşitlik (2.1)'de tanımlı üstel aile formunda

$$f_Y(y; p) = \exp \left\{ y \log \left(\frac{p}{1-p} \right) + n \log(1-p) + \log \binom{n}{y} \right\} \quad (2.9)$$

olarak yazılır. Buna göre, $\theta = \log \left(\frac{p}{1-p} \right)$, $\varphi = 1$, $a(\varphi) = 1$, $b(\theta) = -n \log(1-p)$ ve $c(y, \varphi) = \log \binom{n}{y}$ 'dir.

Poisson Dağılımı: Y yanıt değişkeni Poisson dağılımlı olsun, $Y \sim \text{Poisson}(\lambda)$, $\lambda > 0$. Y 'nin of

$$f_Y(y; \lambda) = \frac{e^{-\lambda} \lambda^y}{y!}, \quad y = 0, 1, 2, \dots \quad (2.10)$$

olup eşitlik (2.1)'de tanımlı üstel aile formunda

$$f_Y(y; \lambda) = \exp \left(y \log \lambda - \lambda - \log(y!) \right) \quad (2.11)$$

olarak yazılır. Buna göre, $\theta = \log(\lambda)$, $\varphi = 1$, $a(\varphi) = 1$, $b(\theta) = e^\theta$ ve $c(y, \varphi) = -\log(y!)$ 'dir.

Normal Dağılım: Y yanıt değişkeni Normal dağılımlı olsun, $Y \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$. Y 'nin olasılık yoğunluk fonksiyonu (oyf)

$$f_Y(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y-\mu)^2}{2\sigma^2} \right\}, \quad -\infty < y < \infty \quad (2.12)$$

olup eşitlik (2.1)'de tanımlı üstel aile formunda

$$f_Y(y; \mu, \sigma^2) = \exp \left\{ \frac{y\mu - \mu^2/2}{\sigma^2} - \frac{1}{2} \left[\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2) \right] \right\} \quad (2.13)$$

olarak yazılır. Buna göre, $\theta = \mu$, $\varphi = \sigma^2$, $a(\varphi) = \varphi$, $b(\theta) = \theta^2/2$,
 $c(y, \varphi) = -\frac{1}{2} \left(\frac{y^2}{\varphi} + \log(2\pi\varphi) \right)$, $b'(\theta) = \mu$ ve $b''(\theta) = 1$ 'dir (Faraway 2006).

Gamma Dağılımı: Y yanıt değişkeni Gamma dağılımlı olsun, $Y \sim \text{Gamma}(\alpha, \beta)$. Y 'nin
oyf

$$f_Y(y; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-\frac{y}{\beta}}, y > 0 \quad (2.14)$$

$E(y) = \mu = \alpha\beta$ olduğundan $\beta = \frac{\mu}{\alpha}$ eşitliği fonksiyonda yerine yazılarak eşitlik (2.14)
düzenlenirse

$$f_Y(y) = \frac{1}{\Gamma(\alpha)} \left(\frac{\alpha}{\mu} \right)^\alpha y^{\alpha-1} e^{-\frac{y\alpha}{\mu}} [\Gamma(\alpha)]^{-1} \alpha^\alpha y^{\alpha-1} \mu^{-\alpha} e^{-\frac{y\alpha}{\mu}} \quad (2.15)$$

olup eşitlik (2.1)'de tanımlı üstel aile formunda

$$f_Y(y) = \exp \left\{ \frac{y/\mu + \log \mu}{-1/\alpha} + \alpha \log(\alpha) + (\alpha - 1) \log(y) - \log(\Gamma(\alpha)) \right\} \quad (2.16)$$

olarak yazılır. Buna göre, $\theta = 1/\mu$, $b(\theta) = -\log(\mu)$, $a(\varphi) = -1/\alpha$,
 $c(y, \theta) = \alpha \log(\alpha) + (\alpha - 1) \log y - \log(\Gamma(\alpha))$, $b'(\theta) = -\frac{1}{\mu}$ ve $b''(\theta) = \frac{1}{\mu^2}$ 'dir (Öztürk
2012).

2.2 Görünüşte İlişkisiz Regresyon Modeli

Yanıt değişkeninin doğrusal ilişkili olduğu çok yanıtlı bir veri kümesinde, yanıt değişkenlerini ayrı ayrı modellemek yerine aralarındaki doğrusal ilişkileri dikkate alarak modelleme yapmak gerekir. Yanıt değişkenleri için oluşturulacak regresyon modellerindeki hata terimleri ilişkiliyse, bu modellere *Görünüşte İlişkisiz Regresyon (Seemingly Unrelated Regression-SUR)* modeli denir (Uzgören 1999). İlk olarak Zellner (1962) tarafından önerilen SUR modeli, hatalar arasındaki ilişkiyi dikkate alan Aitken'in Genelleştirilmiş OLS yöntemi kullanılarak elde edilmiştir. SUR analiziyle elde edilen model tahminlerinin daha etkin olduğu gözlenmiştir (Zellner 1962). Bu nedenle, çok yanıtlı deneylerde SUR analiziyle, yanıtların korelasyon yapısı dikkate alınarak yanıtlar eş anlamlı modellenir.

Çok yanıtlı veri kümesinde yanıt değişkenleri yüksek derecede ilişkili olduğunda, SUR analiziyle elde edilen parametre nokta tahminleri daha düşük standart hataya, parametre aralık tahminleri ise daha dar aralığa sahip olur (Tunçel 2022).

Doğrusal ilişkili r tane yanıt değişkeni için oluşturulacak doğrusal regresyon modeli

$$\mathbf{Y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, 2, \dots, r \quad (2.17)$$

biçimde vektör-matris notasyonuyla tanımlanabilir. Burada, $\mathbf{Y}_j, n \times 1$ boyutlu j . yanıt vektörü; $\mathbf{X}_j, n \times k_j$ boyutlu ve k_j ranklı girdi değişkenlerinin bilinen fonksiyonlarından oluşan tasarım matrisi; $\boldsymbol{\beta}_j, k_j \times 1$ boyutlu bilinmeyen parametreler vektörü ve $\boldsymbol{\varepsilon}_j, n \times 1$ boyutlu j . yanıt vektörüne ilişkin rastgele hata vektörüdür. Burada, r yanıtlı modelin her biri için hataların ilişkisiz, model varyansının sabit ve farklı modeller arasındaki hataların ilişkili olduğu varsayılabilir. Bu durumda tanımlanan varsayımlar

- i. $E(\boldsymbol{\varepsilon}_j) = \mathbf{0}, \quad j = 1, 2, \dots, r$
- ii. $Var(\boldsymbol{\varepsilon}_j) = \sigma_{jj} \mathbf{I}_n, \quad j = 1, 2, \dots, r$

$$\text{iii. } Cov(\boldsymbol{\varepsilon}_j, \boldsymbol{\varepsilon}_l) = \sigma_{jl} \mathbf{I}_n, \quad j, l = 1, 2, \dots, r, \quad j \neq l$$

biçiminde ifade edilebilir. $\boldsymbol{\beta}_j, j = 1, 2, \dots, r$ parametre vektörü için OLS tahmin edicisi

$$\hat{\boldsymbol{\beta}}_j = (\mathbf{X}_j^T \mathbf{X}_j)^{-1} \mathbf{X}_j^T \mathbf{Y}_j \quad (2.18)$$

dir. Eşitlik (2.17)'de j . yanıt değişkeni için verilen doğrusal regresyon modeli genel halde

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.19)$$

eşitliği ile tanımlanabilir. Burada, $\mathbf{Y} = [\mathbf{Y}_1^T : \mathbf{Y}_2^T : \dots : \mathbf{Y}_r^T]^T$, $nr \times 1$ boyutlu yanıt vektörü; $\mathbf{X} = \text{diag}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_r)$, $nr \times p$ boyutlu tasarım matrisi; $\boldsymbol{\beta} = [\boldsymbol{\beta}_1^T : \boldsymbol{\beta}_2^T : \dots : \boldsymbol{\beta}_r^T]^T$, $p \times 1$ boyutlu parametre vektörü ve $\boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_1^T : \boldsymbol{\varepsilon}_2^T : \dots : \boldsymbol{\varepsilon}_r^T]^T$, $nr \times 1$ boyutlu hata vektörüdür. Eşitlik (2.19)'da tanımlı regresyon modeli $E(\boldsymbol{\varepsilon}) = \mathbf{0}_s$ ve $Var(\boldsymbol{\varepsilon}) = \boldsymbol{\psi} = \boldsymbol{\Sigma} \otimes \mathbf{I}_p$ varsayımlarını sağladığında SUR modeli olarak adlandırılır. Burada $\otimes$, Kronecker çarpım operatörü, $\boldsymbol{\Sigma} = (\sigma_{ij}) = \boldsymbol{\Sigma}_d \wedge \boldsymbol{\Sigma}_d : r \times r$ pozitif tanımlı matris, $\boldsymbol{\Sigma}_d = \text{diag}(\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, \dots, \sqrt{\sigma_{pp}}) : r \times r$ pozitif köşegen elemanları olan köşegen matris, $\boldsymbol{\Lambda} = (r_{ij}) : r \times r$ pozitif tanımlı matris ve $s = nr$ 'dir. $\boldsymbol{\beta}$ 'nin en iyi doğrusal yansız tahmin edicisi (Best Linear Unbiased Estimator-BLUE)

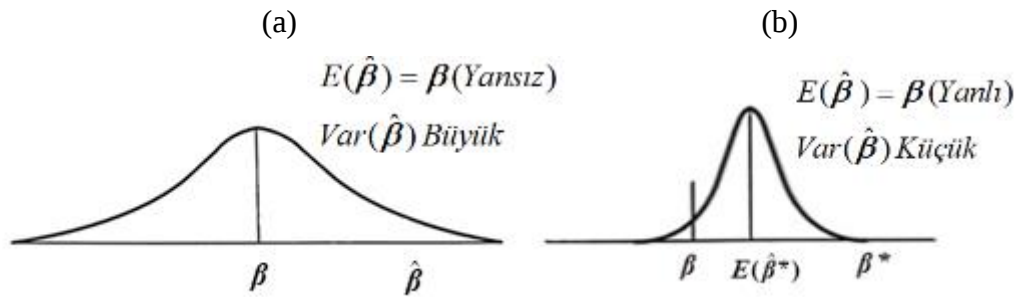
$$\hat{\boldsymbol{\beta}}_{SUR} = (\mathbf{X}^T (\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_n) \mathbf{X})^{-1} \mathbf{X}^T (\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_n) \mathbf{Y} \quad (2.20)$$

biçiminde verilir. Burada, $\boldsymbol{\Sigma}$, yanıtların varyans kovaryans matrisi ve $\mathbf{I}_n$, $n \times n$ boyutlu birim matrisidir (Türkşen vd. 2019). $\hat{\boldsymbol{\beta}}_{SUR}, \boldsymbol{\beta}$ parametre vektörü için SUR tahmin edicisi olur.

2.3 İstatistiksel Daraltıcı Modeller

Çoklu doğrusal regresyon analizinde, OLS yöntemiyle elde edilen parametre tahminlerinin güvenilir olabilmesi, OLS yönteminin temel varsayımlarından biri olan bağımsız değişkenler arasında çoklu bağlantı sorununun olmamasıyla gerçekleşir. Çoklu doğrusal bağlantı durumunda, OLS yönteminin kullanımı parametre nokta tahminlerinin büyük varyansa sahip olmasına neden olur. Çoklu bağlantı sorunu korelasyon matrisinin incelenmesi, varyans artış (şişme) faktörü (Variance Inflation Factor -VIF), öz değerler ve öz vektör analiziyle belirlenir (Montgomery vd. 2012).

Bağımsız değişkenler arasında çoklu doğrusal bağlantı olması durumunda cezalandırılmış doğrusal regresyon modelleri kullanılır ve bu modeller *istatistiksel daraltıcı modeller* olarak adlandırılır. İstatistiksel daraltıcı regresyon yöntemleriyle çoklu bağlantı sorununun boyutu ve hangi bağımsız değişkenlerden kaynaklandığı anlaşılır. Şekil 2.1’de eşitlik (2.19)’da tanımlı β parametre vektörünün yansız ve yanlı tahmin edicilerinin örneklem dağılımı görülmektedir. Şekil 2.1 (a)’da β ’nın OLS yöntemiyle elde edilen tahmin edicisi yansız ve büyük varyanslı iken şekil 2.1 (b)’de β ’nın istatistiksel daraltıcı regresyon yöntemleriyle elde edilen tahmin edicisi yanlı ve küçük varyanslıdır.



Şekil 2.1 β ’nın (a) Yansız ve (b) Yanlı tahmin edicilerinin örneklem dağılımı (Montgomery vd. 2012)

Ridge, LASSO ve Elastik Net regresyonu en çok kullanılan istatistiksel daraltıcı modellerdir. Bu modellerde, ceza terimi kullanılarak β parametre vektörünün en iyi tahmini elde edilir.

2.3.1 Ridge regresyon

Ridge regresyonu, ilk kez Hoerl ve Kennard (1970a, b) tarafından çoklu doğrusal bağlantı sorununu çözen bir yöntem olarak tanımlanmıştır. Ridge regresyon ile regresyon katsayı tahminleri düzenlenerek ya da sıfıra doğru daraltılarak varyanslar önemli ölçüde azaltılabilir. Böylece daha güvenilir parametre tahminleri elde edilir. Ridge regresyonu, L_2 düzenlemesi (regularization) olarak da bilinir.

$$\sum_{i=1}^n \left(Y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \text{ hata kareler toplamı (Sum of Squared Error- SSE) olmak üzere}$$
$$\varphi_R = SSE + \lambda \sum_{j=1}^p \beta_j^2 \quad (2.21)$$

fonksiyonunun minimize edilmesiyle β parametre vektörünün Ridge tahmin edicisi $\hat{\beta}_R$ elde edilir. Burada, $\lambda (\geq 0)$ ayarlanabilir parametre olup $\lambda \sum_{j=1}^p \beta_j^2$ ceza terimi olarak adlandırılır. Eşitlik (2.21)'de, SSE ile $\lambda \sum_{j=1}^p \beta_j^2$ terimlerinin toplamı minimize edilerek en

uygun Ridge regresyon modeli oluşturulur. Ayarlanabilir parametre λ , bu iki terimin regresyon katsayısı tahminleri üzerindeki göreceli etkisini kontrol eder. $\lambda = 0$ olduğunda, ceza teriminin hiçbir etkisi yoktur ve Ridge regresyonu OLS tahminleri üretir ve $\hat{\beta}_R = \hat{\beta}$ olur. Ancak, $\lambda \rightarrow \infty$ olduğunda, ceza terimi büyür ve Ridge regresyon katsayı tahminleri sıfıra yaklaşır. Ridge regresyonu, OLS yönteminin aksine λ 'nın her değeri için farklı bir katsayı tahminleri kümesi üretir. λ için iyi bir değer seçmek kritiktir. En iyi modelin elde edilmesi için kullanılacak λ değeri çapraz doğrulama (Cross Validation-CV) ile belirlenir (James vd. 2023). Bunun için veri seti rastgele bir biçimde eşit büyüklükte k alt örneğe bölünür. $k - 1$ alt örnekleri tahmin modelinin geliştirilmesi için kullanılırken, geri kalan alt örnekler bu modelin doğrulanması için kullanılır. k alt örneklerin her biri sırasıyla doğrulama ve diğerleri model geliştirme için kullanılmak üzere bu işlem k kez gerçekleştirilir. Bir λ değeri aralığı için k ayrı doğrulama sonucu birleştirilerek ve tercih

edilen λ seçilerek genel bir sonuç üretilir. Bu sonuç daha sonra optimal modeli belirlemek için kullanılır.

Eşitlik (2.21)'de tanımlı ceza ve hata kareler toplamlarını minimize edilen Hoerl ve Kennard (1970a) tarafından önerilen Ridge regresyon tahmin edicisi

$$\hat{\beta}_R = [X^T X + K]^{-1} X^T Y \quad (2.22)$$

dir. Burada K , negatif olmayan sabitlerden oluşan bir köşegen matristir, $K = \lambda I_p$, $\lambda \geq 0$ Eşitlik (2.22) düzenlendiğinde

$$\hat{\beta}_R = [X^T X + \lambda I_p]^{-1} X^T Y \quad (2.23)$$

olur. $Z = [I + \lambda(X^T X)^{-1}]$ olarak tanımlandığında, Ridge tahmin edicisi ($\hat{\beta}_R$) ile OLS tahmin edicisi ($\hat{\beta}$) arasındaki ilişki

$$\hat{\beta}_R = Z\hat{\beta} \quad (2.24)$$

biçiminde tanımlanır (Kurtuluş 2001).

2.3.2 LASSO regresyon

LASSO (Least Absolute Shrinkage and Selection Operator - En Az Mutlak Büzülme ve Seçim Operatörü) regresyonu, istatistiksel modelin tahmin doğruluğunu ve yorumlanabilirliğini artırmak için hem değişken seçimini hem de düzenlemeyi sağlayan bir regresyon modelidir. Ridge regresyon modeline alternatif olarak Tibshirani (1996) tarafından önerilen LASSO regresyonda modele katkı sağlayan değişkenler seçilerek modelde tutulur. Bir model oluşturmak için değişkenlere bilinen regresyon yönteminin uygulanması sonucunda modele dahil edilmesi anlamlı olan değişkenler bakımından modelin uyum sağlaması istenir. Bununla birlikte, modelin, aşırı uyum göstermiş olması iyi bir performans olarak değerlendirilmez. Cezalandırma veya düzenleme yaklaşımı, aşırı uyum sorununu çözmek için kullanılır. Bu amaçla regresyon modelleri için bir küçültme ve değişken seçim yöntemi olan LASSO regresyonu uygulanır. LASSO

regresyonuyla modelde tahmin hatasını en aza indiren değişkenler ve bu değişkenlere karşılık gelen regresyon katsayıları belirlenir. Regresyon katsayılarının mutlak değerlerinin toplamını sabit bir değerden (λ) daha az olmaya zorlayıp model parametrelerine bir kısıtlama getirilerek daraltma uygulanır. Bu uygulama sonrasında regresyon katsayısı sıfır olan değişkenler modelden çıkarılır. λ 'nın seçimi, Ridge regresyondaki uygulamasına benzer olarak k -kat CV yöntemi kullanılarak yapılır (Ranstam ve Cook 2018).

L_1 düzenlemesi olarak da bilinen LASSO regresyonu

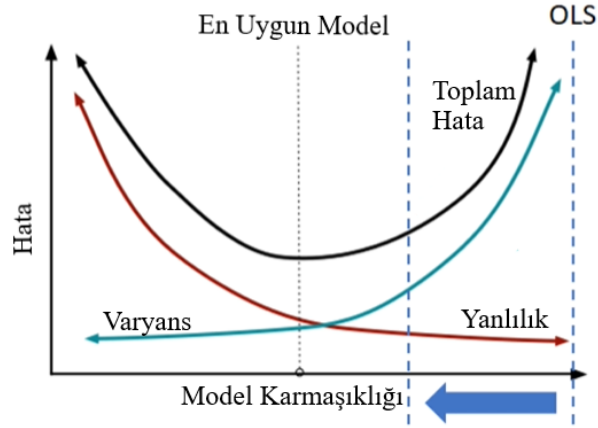
$$\varphi_L = SSE + \lambda \sum_{j=1}^p |\beta_j| \quad (2.25)$$

fonksiyonunun minimize edilmesiyle β parametre vektörünün LASSO tahmin edicisi $\hat{\beta}_L$ elde edilir (James vd. 2023). Burada, SSE, hata kareler toplamı, λ ayarlanabilir

parametre ve $\lambda \sum_{j=1}^p |\beta_j|$ ceza terimidir.

Eşitlik (2.21)'de ve eşitlik (2.25)'te görüldüğü gibi Ridge ve LASSO regresyonu benzer minimizasyon fonksiyonlarına sahiptir. Fakat, tek farklılık ceza terimi fonksiyonunun tanımlanma biçimidir.

LASSO regresyonda modele katkısı olmayan değişkenlerin katsayısı sıfıra eşitlendiğinden bu yöntem değişken seçimi içinde uygulanan bir yöntemdir. Böylece, modele katkı sağlayan değişkenlerin seçilmesiyle modelin yorumlanması kolaylaşır. Ridge regresyon yönteminde tüm değişkenler her zaman modelde yer alırken LASSO regresyon yönteminde katsayısı sıfır olan değişkenler modelde yer almaz. Şekil 2.2'de varyans ve yanlılığın değişimi görülmektedir.



Şekil 2.2 Varyans/Yanlılık ödünleşimi (Pichler ve Hartig 2022)

2.3.3 Elastik Net regresyonu

Elastik Net, LASSO ve Ridge regresyon modellerinin L_1 ve L_2 düzenlemelerini doğrusal olarak birleştiren bir regresyon modelidir. Zou ve Hastie (2005) tarafından önerilen Elastik Net regresyonu, LASSO regresyona benzer biçimde eş anlı olarak değişken seçimi ve regresyon modelleri için katsayılar üzerinde daraltma yapar. Elastik Net, LASSO'nun bir uzantısıdır ve bağımsız değişkenler arasındaki yüksek korelasyonlara karşı dayanıklıdır (Ogutu vd. 2012). Elastik Net regresyonu L_1 düzenlemesiyle modele katkısı olmayan değişkenlerin katsayılarını sıfırlayarak değişken seçimi yaparken L_2 düzenlemesiyle değişkenlerin katsayılarını sınırlayarak aşırı uyumu önler. Elastik Net, L_1 ve L_2 düzenlemelerini birlikte kullandığı için esnektir. Bir veri setinde korelasyonu yüksek olan değişkenler aralarında değişken grubu oluşturur. LASSO regresyonu birbiriyle ilişkili değişken gruplarının varlığında yalnızca bir değişken seçme eğilimindeyken hangi değişkenin seçileceğini önemsemez. Fakat, Elastik Net regresyonu ile LASSO regresyonun aksine birbiriyle ilişkili bu değişken grupları seçilebilir. Elastik Net regresyonu, tahmin doğruluğu bakımından da LASSO regresyondan daha iyi bir performans gösterir. Değişken sayısı, gözlem sayısından büyük olduğunda Elastik Net regresyonunun kullanımı daha avantajlıdır.

Elastik Net regresyonda

$$\varphi_{\text{EN}} = \text{SSE} + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2 \quad (2.26)$$

fonksiyonunun minimize edilmesiyle β parametre vektörünün Elastik Net tahmin edicisi

$\hat{\beta}_{\text{EN}}$, $\alpha = \lambda_2 / (\lambda_1 + \lambda_2)$ olmak üzere, $\sum_{j=1}^p |\beta_j| = |\beta|_{L_1}$ ve $\sum_{j=1}^p \beta_j^2 = |\beta|_{L_2}$ olarak

tanımlandığında $(1-\alpha)|\beta|_{L_1} + \alpha|\beta|_{L_2} \leq t$ koşulu altında $\hat{\beta}_{\text{EN}}$ tahmin edicisi bulunur.

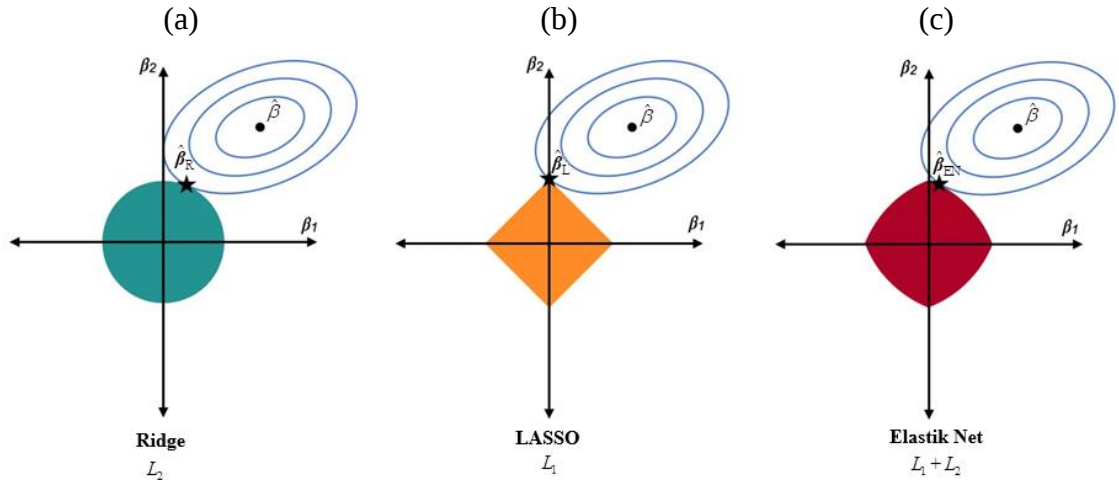
Burada $(1-\alpha)|\beta|_{L_1} + \alpha|\beta|_{L_2}$ toplamı Elastik Net tahmin edicisinin ceza terimidir. Elastik

Net modeli, $\alpha = 1$ iken Ridge, $\alpha = 0$ iken LASSO regresyon modeline dönüşür. $\alpha \in [0,1]$

olduğunda Elastik Net, Ridge ve LASSO regresyonun özelliklerine sahip olur. Elastik

Net regresyonda, λ ve α ayarlanabilir parametreleri seçimi Ridge ve LASSO regresyonda yapıldığı gibi CV yöntemiyle belirlenir (Vergili 2023).

β_1 ve β_2 parametrelerini içeren bir regresyon modeli için Ridge, LASSO ve Elastik Net tahmin edicileri şekil 2.3'te verilmiştir.



Şekil 2.3 (a) Ridge, (b) LASSO ve (c) Elastik Net regresyon modelleri için tahmin ediciler (James vd. 2023)

Şekil 2.3 (a)-(c)'de taralı bölgeler sırasıyla, Ridge, LASSO ve Elastik Net regresyon modellerinin kısıt bölgelerini ifade eder. $\hat{\beta}$, OLS tahmin edicisi olmak üzere, şekil 2.3

(a)'da Ridge tahmin edicisi $\hat{\beta}_R$ 'nin sıfırdan farklı parametre tahminlerini içerdği görülmektedir. LASSO tahmin edicisi $\hat{\beta}_L$ 'de ise, bir parametre tahmin değerinin sıfır olduğu şekil 2.3 (b)'de açıktır. Şekil 2.3 (c)'den ise Elastik Net tahmin edicisi $\hat{\beta}_{EN}$ 'nin, $\hat{\beta}_R$ ve $\hat{\beta}_L$ arasında değerler aldığı söylenir. $\hat{\beta}_R$, $\hat{\beta}_L$ ve $\hat{\beta}_{EN}$ tahmin edicileri, SSE konturlarının kısıt bölgesine ilk kez temas ettiği noktaya karşılık gelir.

Bu bölümde çok yanıtlı karma verili problemlerin modellenmesinde yanıt değişkenlerinin dağılımı, yanıt değişkenleri arasında doğrusal ilişki olup olmaması, girdi değişkenlerinin çoklu bağlantı sorunu içerip içermemesi gibi durumların dikkate alınması gerektiğine yer verilmiştir. Her bir durumda uygulanacak modellere şematik olarak çizelge 2.4'te yer verilmiştir.

Çizelge 2.4 Farklı durumlarda uygulanan modelleme yöntemleri

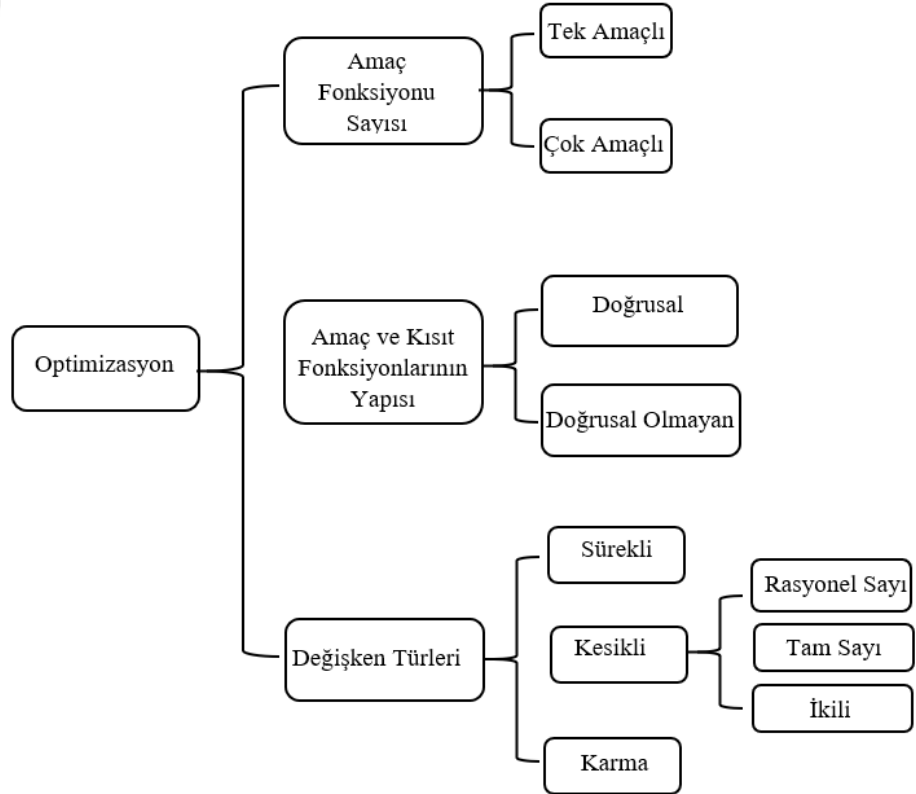
Durumlar	Modelleme Yöntemleri
Yanıt değişkeni Normal dağılımlı olmadığı	GLM
Girdi değişkenleri arasında doğrusal ilişki olduğunda	Ridge LASSO Elastik Net
Yanıt değişkenleri arasında doğrusal ilişki olduğunda	SUR

3. OPTİMİZASYON MODELLERİNİN SINIFLANDIRILMASI

Optimizasyon, bir sistemde, belirli kısıtlar altında belirlenmiş bir amaç fonksiyon değerinin problemin türüne göre maksimize ya da minimize edilmesi amacıyla karar değişkenlerinin alacağı değerlerin belirlenmesidir. Başka bir deyişle, istenen çıktıyı elde etmek için sistemin girdi değişkenlerinin değerlerinin ne olacağının belirlenmesidir (Küçükkoç 2020).

Optimizasyon modeli, probleme ilişkin amacın/amaçların ve kısıtın/kısıtların matematiksel olarak ifade edilmesi olup optimal çözüm elde ederek karar verme konusunda optimizasyon sürecinin önemli bir bileşenidir (Türkşen 2023).

Optimizasyon modellerinin sınıflandırılması, amaç fonksiyonunun sayısına, amaç ve kısıt fonksiyonlarının yapısına ve değişken türlerine göre yapılabilir (Yang 2010). Şekil 3.1’de türlerine göre optimizasyon modellerinin sınıflandırılması verilmiştir.



Şekil 3.1 Optimizasyon modellerinin sınıflandırılması

3.1 Amaç Fonksiyonu Sayısına Göre Optimizasyon

Optimizasyon modelleri, minimum ya da maksimum değeri elde edilmek istenilen amaç fonksiyonlarının sayısına göre tek amaçlı optimizasyon ve çok amaçlı optimizasyon (MOO) olmak üzere iki başlıkta değerlendirilir.

3.1.1 Tek amaçlı optimizasyon

Tek amaçlı optimizasyon modeli genel olarak

$$\begin{aligned} & \min / \max f(\mathbf{X}) \\ & h_i(\mathbf{X}) = 0 \quad , \quad i = 1, 2, \dots, k \\ & g_j(\mathbf{X}) \{ \leq, \geq \} 0 \quad , \quad j = 1, 2, \dots, l \\ & \mathbf{X} \in \mathbb{R} \end{aligned} \quad (3.1)$$

biçiminde tanımlanabilir. Burada, $\mathbf{X} = [X_1 \ X_2 \ \dots \ X_p]$ karar değişkenleri vektörü, f amaç fonksiyonu, $h_i, i = 1, 2, \dots, k$ eşitlik ve $g_j, j = 1, 2, \dots, l$ eşitsizlik biçiminde tanımlı kısıt fonksiyonlarıdır. $\mathbf{X}$ karar değişkenleri vektörünün aldığı değerleri gösteren vektör $\mathbf{x}$ olmak üzere, (3.1) ifadesiyle tanımlı model

$$\begin{aligned} & \min / \max f(\mathbf{x}) \\ & h_i(\mathbf{x}) = 0 \quad , \quad i = 1, 2, \dots, k \\ & g_j(\mathbf{x}) \{ \leq, \geq \} 0 \quad , \quad j = 1, 2, \dots, l \\ & \mathbf{x} \in \mathbb{R} \end{aligned} \quad (3.2)$$

biçiminde de ifade edilebilir (Türkşen 2023).

(3.2) ifadesiyle tanımlı modelde, tüm karar değişkenleri sürekli değerler alması durumunda model *sürekli değerli optimizasyon modeli*, tüm karar değişkenlerinin tam sayılı değerler alması durumunda model *saf tam sayılı optimizasyon modeli*, bazı karar değişkenlerinin sürekli, bazılarının tam sayı değerleri alması durumunda model *karma tam sayılı optimizasyon modeli* ve bazı karar değişkenlerinin sürekli, bazılarının kesikli değerler alması durumunda model *karma kesikli değerli optimizasyon modeli* olarak adlandırılır.

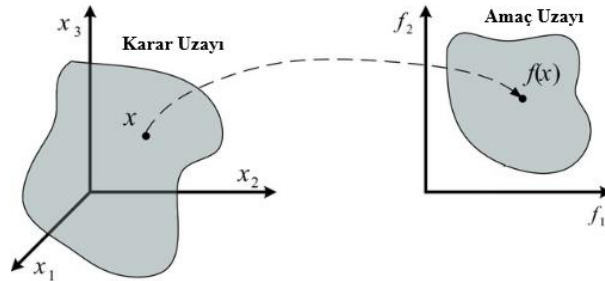
3.1.2 Çok amaçlı optimizasyon

MOO probleminin matematiksel modeli

$$\begin{aligned} \min / \max \quad & f(\mathbf{x}) = [f_1(\mathbf{x}) \ f_2(\mathbf{x}) \ \dots \ f_M(\mathbf{x})] \\ & h_i(\mathbf{x}) = 0 \quad , \quad i = 1, 2, \dots, k \\ & g_j(\mathbf{x}) \{ \leq, \geq \} 0 \quad , \quad j = 1, 2, \dots, l \\ & \mathbf{x} \in \mathbb{R} \end{aligned} \quad (3.3)$$

biçiminde tanımlanır. Burada, $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]$ biçiminde tanımlı karar değişken değerleri vektörü, $\mathbf{f} = [f_1 \ f_2 \ \dots \ f_M]$ amaç fonksiyonu vektörü, $h_i, i = 1, 2, \dots, k$ eşitlik ve $g_j, j = 1, 2, \dots, l$ eşitsizlik biçiminde tanımlı kısıt fonksiyonlarıdır. Amaç ve kısıt fonksiyonlarının tümü doğrusal ise model, *çok amaçlı doğrusal optimizasyon modeli*, amaç veya kısıt fonksiyonlarından en az bir tanesi doğrusal değil ise model, *çok amaçlı doğrusal olmayan optimizasyon modeli* olur.

MOO'da tek amaçlı fonksiyondan farklı olarak amaç fonksiyonları karar değişkenleri uzayına ek olarak çok boyutlu bir uzay oluşturmaktadır. Bu uzay, *amaç uzayı* olarak adlandırılır. Karar değişkenleri uzayındaki her x çözümü için, amaç uzayında $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_M]^T$ ile gösterilen bir nokta vardır. n boyutlu bir çözüm vektörüyle M boyutlu amaç vektörü eşleştirilir. Şekil 3.2'de karar uzayı ve karar uzayına karşılık gelen amaç uzayı arasındaki ilişki gösterilmiştir.



Şekil 3.2 Karar uzayı ve karar uzayına karşılık gelen amaç uzayı arasındaki ilişki (Deb 2001)

(3.3) ifadesiyle tanımlı MOO probleminin optimizasyonu sonucunda tek bir optimal çözümden söz edilemez. Elde edilen hiçbir çözüm bütün amaç fonksiyonlarını aynı anda optimize edemeyeceğinden ideal çözüme ulaşamaz. MOO problemlerinin çözümü Pareto çözümler ile karakterize edilir. Matematiksel olarak her optimal çözüm MOO probleminin aynı derecede kabul edilebilir bir çözümdür. Bir Pareto çözüm kümesi birbirine ne baskın ne de basılgin olan çok sayıda alternatif çözümden oluşur (Türkşen 2011).

Pareto çözüm kümesi için baskınlık kavramının tanımlanmasında amaç fonksiyonlarının minimum ya da maksimum yapılması etkilidir. Minimum değerlerinin elde edilmesi istenilen MOO problemi için baskınlık kavramı aşağıdaki gibi tanımlanır.

Tanım 3.1 Baskınlık: $\mathbf{x}^1, \mathbf{x}^2 \in S$ olmak üzere $\mathbf{x}^1$ çözümü, $\mathbf{x}^2$ çözümüne

1. $\mathbf{x}^1$ çözümü, $\mathbf{x}^2$ çözümünden tüm amaç fonksiyonları üzerinde daha kötü değilse
 $(f_j(\mathbf{x}^1) \leq f_j(\mathbf{x}^2), \forall j = 1, 2, \dots, M)$
2. $\mathbf{x}^1$ çözümü, en az bir amaç fonksiyonunda $\mathbf{x}^2$ çözümünden mutlaka daha iyiye
 $(f_j(\mathbf{x}^1) < f_j(\mathbf{x}^2), \exists j = 1, 2, \dots, M)$

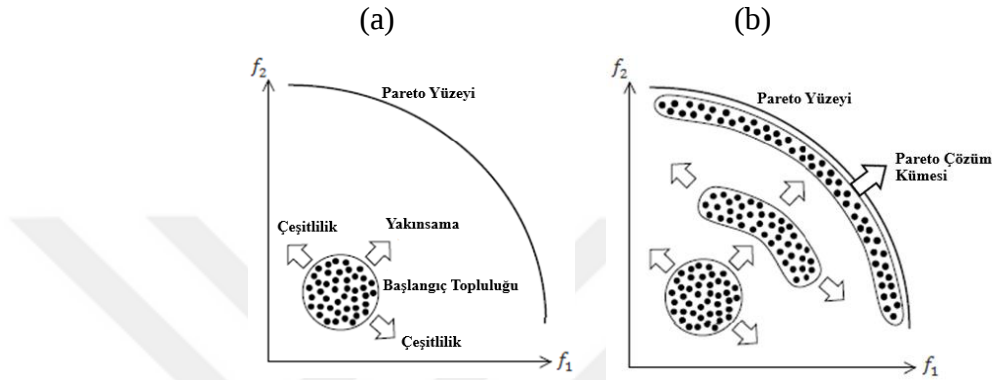
her iki durum sağlandığında baskın olur. Aksi halde verilen koşullardan birinin ihlali durumunda $\mathbf{x}^1$ çözümü, $\mathbf{x}^2$ çözümüne baskın değildir. Eğer, $\mathbf{x}^1$ çözümü, $\mathbf{x}^2$ çözümüne baskın ($\mathbf{x}^1 \prec \mathbf{x}^2$) ise

- $\mathbf{x}^2$, $\mathbf{x}^1$ çözümü tarafından bastırılmıştır,
- $\mathbf{x}^1$, $\mathbf{x}^2$ çözümü tarafından bastırılamaz ya da
- $\mathbf{x}^1$, $\mathbf{x}^2$ 'den daha kötü bir çözüm değildir.

Tanım 3.2 Pareto-optimallik: Eğer $f_j(\mathbf{x}^*) \leq f_j(\mathbf{x}), \forall j = 1, 2, \dots, M$ ve en az bir p indisi için $f_p(\mathbf{x}^*) \leq f_p(\mathbf{x})$ olacak şekilde bir başka karar vektörü $\mathbf{x} \in S$ yoksa $\mathbf{x}^* \in S$ Pareto optimaldir (Miettinen 1998). Pareto optimal küme, karar vektörlerinden oluşan bir

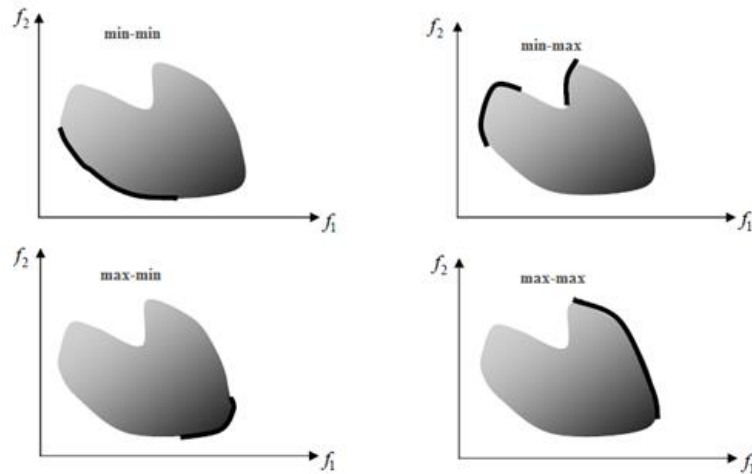
kümedir. Pareto optimal yüzey ise tüm amaç fonksiyonları göz önüne alındığında, baskın olan tüm karar vektörlerinden oluşan yüzeydir.

MOO problemlerini çözmek için kullanılan algoritmalar yakınsama ve çeşitlilik hedeflerine odaklanır (Deb 2001). Bu iki hedef, iki amaç fonksiyonu maksimum yapılması istenilen bir MOO problemi için şekil 3.3'te temsil edilmektedir.



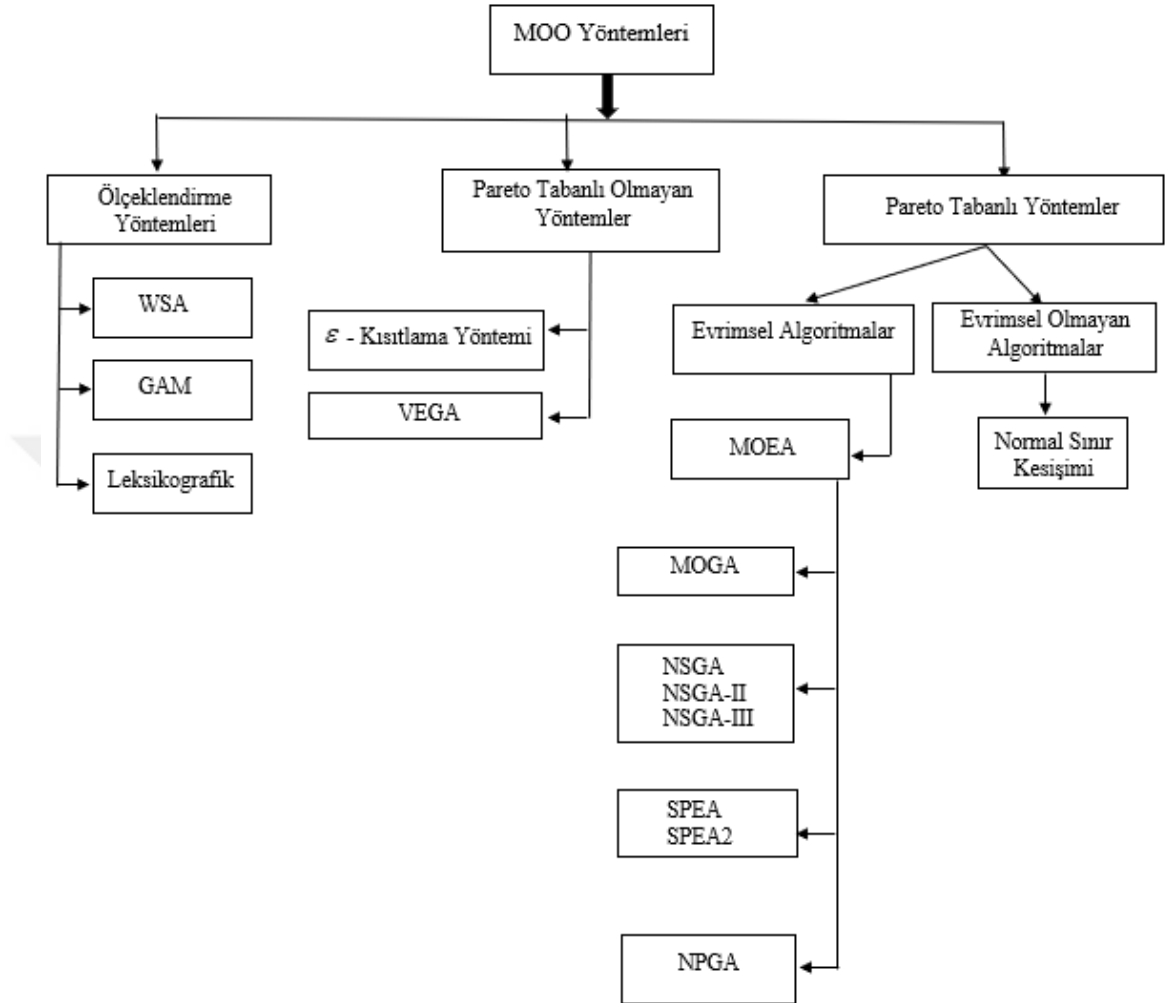
Şekil 3.3 (a) Başlangıç topluluğu ve Pareto yüzeyi, b) Pareto çözüm kümesi (Ishibuchi ve Shibata 2004)

İlk hedef olan yakınsama, yakınsanan Pareto kümesinin Pareto çözüm kümesine yakınlığını temsil eder. İkinci hedef olan çeşitlilik ise yakınsanan Pareto kümesinin hem yayılımını hem de dağılımını temsil eder. Şekil 3.4'te, iki amaç fonksiyonu için oluşabilecek dört farklı durumda (*min-min*, *min-max*, *max-min*, *max-max*) elde edilen Pareto çözüm kümeleri gösterilmektedir.



Şekil 3.4 İki amaç fonksiyonu için olabilecek dört farklı durumda elde edilen Pareto çözüm kümeleri (Türkşen 2011)

MOO problemlerini çözmek için literatürde yer alan çok sayıda yöntem vardır. Şekil 3.5'te kullanılan başlıca MOO yöntemlerinin sınıflandırılması yer almaktadır.



Şekil 3.5 MOO yöntemlerinin sınıflandırılması (Gambier 2008'den değiştirilerek alınmıştır)

Şekil 3.5'te belirtilen Ölçeklendirme Yöntemleri, birden fazla amaç fonksiyonunun tek bir amaç fonksiyonuna dönüştürülerek optimizasyonun gerçekleştirildiği yöntemlerdir (Rainer 2012). Bu başlıkta en çok tercih edilen yöntemler, Toplam Ağırlıklandırma Yaklaşımı (Weighted Sum Approach - WSA), Hedef Atama Yöntemi (Goal Attainment Method - GAM) ve Leksikografik Yöntem'dir.

Pareto Tabanlı Olmayan Yöntemlerden olan ε -Kısıtlama Yöntemi'nde çok sayıdaki amaç fonksiyonlarından herhangi biri amaç fonksiyonu olarak ele alınıp diğer amaç

fonksiyonları kısıt fonksiyonları olarak değerlendirilir. Vektör hesaplamalı GA (Vector Evaluated Genetic Algorithm-VEGA), bir sonraki popülasyonun alt kümesinin tek bir amaç tarafından seçilmesi ve gelecek popülasyonun oluşturulması için tüm alt kümelerin birlikte kullanılması prensibini içerir. Böylece gelecek popülasyonun her amacın en iyi bireylerinden oluşması sağlanır (Kaya ve Fırlalı 2016).

Pareto Tabanlı Yöntemler’de, amaç fonksiyonlarının eş anlı optimizasyonu yapılır. Evrimsel Algoritmalar (Evolutionary Algorithms-EA), Pareto tabanlı yöntemler arasında en çok tercih edilenidir. Çok amaçlı EA (Multi Objective Evolutionary Algorithm-MOEA) popülasyon tabanlı yapay zeka optimizasyonu algoritmaları içerir. Örneğin, Parçacık Sürü Optimizasyonu, Karınca Kolonisi Algoritması, Diferansiyel Evrim Algoritması vb. MOO problemlerini çözmek için önerilen EA’lardan en çok ilgi görenleri Fonseca ve Fleming’in (1993) önerdiği Çok Amaçlı GA (Multi Objective Genetic Algorithm - MOGA), Srinivas ve Deb’in (1995) önerdiği Baskın Sıralı GA (Nondominated Sorting Genetic Algorithm - NSGA), Horn vd.’nin (1994) önerdiği Niş Pareto GA (Niche Pareto Genetic Algorithm-NPGA) ve Zitzler vd. (2000) önerdiği Güçlü Pareto EA (Strength Pareto Evolutionary Algorithm-SPEA)’dır. Bu algoritmalar, basit bir EA’yı bir MOEA’ya dönüştürmek için gerekli ek operatörler içerir. Operatördeki ortak özellikler

- i. Baskın sıralamaya dayalı olarak popülasyon üyelerine baskınlık sırası atamak
 - ii. Aynı baskın yüzeyin çözümleri arasındaki çözüm çeşitliliğini korumak
- tır.

3.2 Fonksiyon Yapılarına Göre Optimizasyon

Optimizasyon modelleri, amaç ve/veya kısıt fonksiyonlarının doğrusal olup olmamasına göre doğrusal ve doğrusal olmayan optimizasyon başlıklarında incelenebilir. Literatürde genel olarak, doğrusal optimizasyon Doğrusal Programlama (Linear Programming- LP) ve doğrusal olmayan optimizasyon ise Doğrusal Olmayan Programlama (Nonlinear Programming-NLP) olarak adlandırılır.

3.2.1 Doğrusal optimizasyon

Optimizasyon problemlerinin çözümünde bir araç olan LP, doğrusal kısıtlar altında doğrusal bir fonksiyonun optimizasyonunu içeren özel bir matematik programlama türüdür. Bir optimizasyon modelinde amaç ve kısıt fonksiyonları doğrusal ve karar değişkenleri sürekli değişkenlerden oluşuyor ise LP modeli olarak tanımlanır. LP modeli en sık kullanılan optimizasyon modelidir (Lorcu 2016, Örkücü vd. 2024). LP modelinin oluşturulabilmesi için toplamsallık, oransallık, belirlilik ve bölünebilirlik varsayımlarının sağlanmış olması gerekir (Türkşen 2023).

LP modellerinin matematiksel formülü (3.2) ifadesiyle aynıdır. LP problemlerinin çözümünde sıklıkla kullanılan yöntemler, Grafik Yöntemi ve Simpleks Yöntemi'dir.

3.2.2 Doğrusal olmayan optimizasyon

Çoğunlukla NLP modelleriyle mühendislik tasarım optimizasyonunda karşılaşılır ve karar/tasarım değişkenleri sürekli değerli olarak ele alınır. Optimizasyon problemlerinin çoğu kısıtlı formdadır. NLP problemlerinin çözümü için geliştirilmiş çeşitli yöntemler mevcuttur. Değişken sayısının arttığında ve veri tipi değiştiğinde problemlerin zorluk derecesi artabilir. Böyle problemlerin deterministik yöntemlerle çözümü, hem problemin yapısına bağlı olarak modellemede hem de çözüm sürecinde zorluklar içerir. Bu problemlerin üstesinden gelebilmek için sezgisel yöntemler geliştirilmiştir. Özellikle popülasyon temelli sezgisel yöntemler, çok noktalı arama yaklaşımları sayesinde kısa sürede sonuç vermektedir (Keskintürk 2006). NLP problemlerinin çözümünde sıklıkla kullanılan yöntemler, Kuhn-Tucker Yöntemi, Lagrange Çarpanları Yöntemi, Quasi-Newton Algoritması, Nelder-Mead Simpleks Algoritması, Trust Region Algoritması, Interior Point Algoritması ve Sezgisel Yöntemler'dir (Rao 2020).

3.3 Değişken Türlerine Göre Optimizasyon

Gerçek dünya problemlerine ilişkin oluşturulan optimizasyon modellerindeki karar değişkenlerinin aldığı değerler optimizasyon sürecinde oldukça önemlidir. Optimizasyon modelleri karar değişkenlerinin türlerine göre sürekli, kesikli ve karma optimizasyon modelleri olarak üç grupta incelenebilir.

3.3.1 Sürekli optimizasyon

Karar değişkenlerinin sürekli olarak tanımlandığı problemler, sürekli optimizasyon problemleri olarak adlandırılır. (3.2) ifadesinde tanımlı matematiksel model, sürekli optimizasyon modeli olarak dikkate alınabilir. Sürekli optimizasyon modelindeki $\mathbb{R}$, sürekli değişkenlerin aldığı değerlerin reel sayılar kümesinde tanımlı olduğunu belirtir. Sürekli değişkenlerin tanım aralığı $\mathbb{R}$ 'de, $[a,b]$ gibi belirli bir aralık olarak da tanımlanabilir. Sürekli optimizasyon problemlerini, türeve dayalı optimizasyon yöntemleriyle örneğin, Newton Yöntemi, En Dik İniş Yöntemi, Levenberg-Marquardt Yöntemi ve türevden bağımsız optimizasyon yöntemleriyle örneğin, Nelder-Mead Simpleks Yöntemi ve Sezgisel Yöntemler kullanarak çözmek mümkündür.

3.3.2 Kesikli optimizasyon

Karar değişkenlerinin kesikli olarak tanımlandığı problemler, kesikli optimizasyon problemleri olarak adlandırılır (Bakır ve Altunkaynak 2003). Kesikli değişkenler kümesi tam sayılı, ondalıklı sayılı ve ikili değerler alan değişkenler kümesini kapsar. Bu nedenle, hem tam sayılı hem de ondalıklı sayılı değişkenleri ifade etmek için *kesikli değişkenler* terimi kullanılır. Kesikli optimizasyon modeli (3.2) ifadesine benzer olarak

$$\begin{aligned} \min / \max f(\mathbf{x}) \\ h_i(\mathbf{x}) &= 0, \quad i = 1, 2, \dots, k \\ g_j(\mathbf{x}) &\{\leq, \geq\} 0, \quad j = 1, 2, \dots, l \\ \mathbf{x} &\in \mathbb{N} \cup \mathbb{Z} \cup \mathbb{Q} \end{aligned} \quad (3.4)$$

biçiminde tanımlanır. Burada, $\mathbb{N}$, $\mathbb{Z}$ ve $\mathbb{Q}$, karar değişkenlerinin aldığı değerler kümesi olarak sırasıyla doğal sayı, tam sayı ve rasyonel sayı kümeleridir.

Kesikli optimizasyon problemlerini çözmek için bilinen yöntemlerden en çok kullanılanlardan biri Dal-Sınır (Branch-Bound) Yöntemidir. Dal-Sınır Yöntemi'nde, sistematik olarak tüm numaralandırma azaltılmaya çalışıldığından bu yöntem *örtük numaralandırma yöntemi (implicit enumeration method)* olarak da adlandırılır (Arora ve Huang 1994).

Kesikli optimizasyon problemleri, karar değişkenlerinin tamamının ya da bir kısmının tam sayılı olması durumunda *tam sayılı programlama problemleri* olarak tanımlanır. Gerçek hayatta karşılaşılan optimizasyon problemlerinin modellenmesinde ilgilenilen karar değişkenleri tam sayılı olabilir. Örneğin, üretim problemleri, atama problemleri vb. Karar değişkenlerinin tamamının ya da bir kısmının tam sayılı olduğu matematiksel programlama türüne *Tam Sayılı Programlama (Integer Programming-IP)* denir. Tüm karar değişkenlerinin 0 ya da 1 değerlerinden oluşan IP problemine, *0-1 Tam Sayılı Programlama* problemi denir. IP'nin bir kısmı zor optimizasyon problemleri sınıfında yer alır. Çoğu IP problemleri, NP (Non-Polynomial) sınıfından problemlerdir ve üstel hesaplama gerektirir (Bakır ve Altunkaynak 2003). IP problemlerinin çözümünde, problemin LP modeli olarak çözümünün yapılması ve optimal çözümde karar değişkenlerinin aldığı değerleri en yakın tam sayı değerlerine yuvarlama yolunun izlenmesi genel bir yaklaşım olarak düşünülebilir. Ancak bu yöntemde, yuvarlanan yaklaşık çözüm değerlerinin, problem kısıtlarını sağlama garantisi yoktur. Özellikle büyük boyutlu problemlerde bu yöntem optimallikten ve problem kısıtlarını sağlamaktan oldukça uzak çözüm sonuçları doğurabilmektedir. Yuvarlama yönteminin bu dezavantajlarından dolayı özel algoritmalar geliştirilmiştir (Ayan vd. 2021). Kesikli optimizasyon problemlerinin çözümünde uygulanan Dal-Sınır Yöntemi, IP problemlerinin çözümünde de sıklıkla kullanılır.

Dal-Sınır Yöntemi'nde, tüm mümkün çözüm kümeleri arasından optimal çözümü bulmak için

- Dallandırma/Ayrıştırma,
- Sınırlama ve
- Bağlama/Düğümleme

işlemleri uygulanır. Geniş bir arama uzayında arama yapmak yerine daha küçük alt arama uzaylarında arama yapmanın kolay ve hızlı olması nedeniyle ana problem önce alt problemlere ayrılarak *dallandırma* yapılır. Dallandırma yapılan alt problemlerin çözümleriyle ana problemin sınırları belirlenir. Tüm karar değişkenlerinin değerleri tam sayılı olana kadar karar değişkenlerinin değerlerine ilişkin yeni kısıtlar modele eklenerek dallandırma ve sonlandırma işlemleri yapılır. Karar değişkenleri tam sayı olan tüm mümkün çözümler, *aday çözüm* olarak adlandırılır. Her bir aşamada aday çözüm alt kümelerinden bir tanesi seçilir ve bu küme daha alt çözüm kümelerine ayrıştırılır. Eğer alt küme boş (karşılık gelen aday çözüm uygun değil) ise, bu aday çözüm '*bağlandı*' (*fathmoded*) denir. Bu dal üzerinden daha ileri gidilemez. Çünkü, bu yöndeki ilerlemeler uygun çözümler vermeyecektir. Eğer aday çözüm bağlanmamışsa, bu alt çözüm kümesi daha basit alt kümelere dallandırma kısıtları ile optimum çözüm elde edilinceye kadar ayrıştırma işlemi devam eder. Dal- Sınır Yöntemi'nin algoritmik adımları aşağıda verilmiştir.

Adım 1: Tüm tam sayı ya da değişkenler üzerindeki 0-1 kısıtları ihmal edilerek (LP gevşetmesi) değişkenler sürekli değişken gibi problem çözülür.

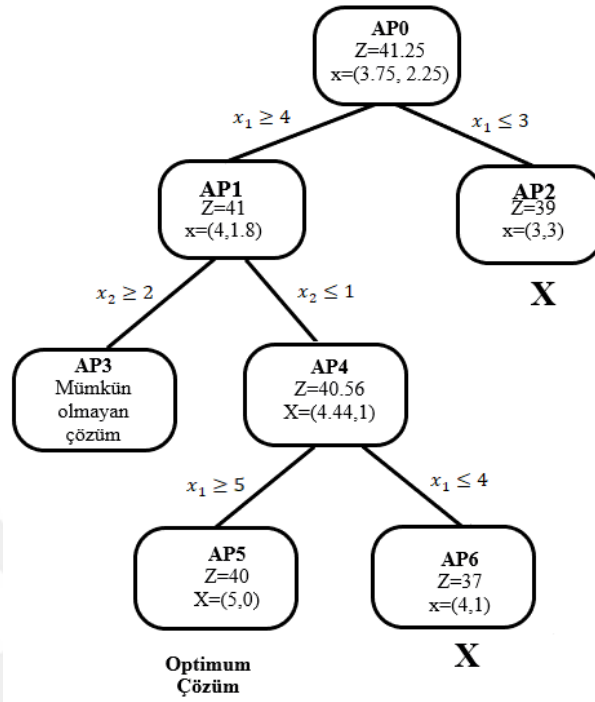
Adım 2: IP'nin optimal çözümünü elde etmek için LP gevşetmesinin uygun çözüm bölgesi dallandırma stratejisi uygulanarak alt problemlere ayrıştırılır.

Adım 3: Alt problemler üzerine tam sayı kısıtları eklenir. Uygun olmayan çözümler varsa dal budanır.

Adım 4: Tüm karar değişkenlerinin değerleri tam sayılı olana kadar yeni kısıtlar modele eklenerek dallandırma işlemi devam eder.

Adım 5: Eğer bir düğümde tam sayılı çözüm veya mümkün olmayan çözüm elde edilmişse bu düğümden dallandırma yapılmaz ve bu düğüm bağlanır.

Şekil 3.6’da bir optimum arama ağacı örneği verilmiştir.



Şekil 3.6 Optimum arama ağacı örneği (Bakır ve Altunkaynak 2003)

3.3.3 Karma optimizasyon

Karar değişkenlerinin sürekli ve kesikli olarak tanımlandığı problemler, *karma optimizasyon problemleri* olarak adlandırılır. Karma optimizasyon modeli

$$\begin{aligned}
 & \min / \max f(x) \\
 & h_i(x) = 0 \quad , \quad i = 1, 2, \dots, k \\
 & g_j(x) \{ \leq, \geq \} 0 \quad , \quad j = 1, 2, \dots, l \\
 & x \in \mathbb{R} \cup \mathbb{N} \cup \mathbb{Z} \cup \mathbb{Q}
 \end{aligned} \tag{3.5}$$

biçiminde ifade edilebilir. Burada karar değişkenlerinin aldığı değerler kümesi reel sayı, doğal sayı, tam sayı ve rasyonel sayı olabilir.

(3.5) ifadesiyle tanımlı karma optimizasyon modelinde amaç fonksiyonu ve kısıt fonksiyonlarının tamamı doğrusal ise model LP modeli olur. Karar değişkenlerinin karma verili değer alması nedeniyle de (3.5) modeli *Karma Tam Sayılı Doğrusal Programlama*

(*Mixed Integer Linear Programming* - MILP) modeli olarak adlandırılır. MILP problemlerini çözmek için kullanılan yöntemler, Yuvarlama Yöntemi, Kesen Yüzeyler Yöntemi, Dal-Sınır Yöntemi ve Sezgisel Yöntemlerdir.

(3.5) ifadesiyle tanımlı karma optimizasyon modelinde amaç fonksiyonu ve kısıt fonksiyonlarından en az biri doğrusal olmayan fonksiyon ise model NLP olur. Karar değişkenlerinin karma verili değer alması nedeniyle de (3.5) modeli *Karma Kesikli Doğrusal Olmayan Programlama (Mixed Discrete NonLinear Programming* - MDNLP) modeli olarak adlandırılır. (3.5) ifadesiyle tanımlı karma optimizasyon modelinde $x \in \mathbb{R} \cup \mathbb{N} \cup \mathbb{Z}$ ise model *Karma Tam Sayılı Doğrusal Olmayan Programlama (Mixed Integer NonLinear Programming* - MINLP) modeli olarak adlandırılır.

MINLP ve MDNLP problemlerini çözmek için birçok algoritma geliştirilmiştir. Bu tür problemler sıklıkla endüstri, finans, mühendislik tasarımı, yönetim bilimi, süreç akış şemaları, portföy seçimi, gaz ve su dağıtım şebekelerinin optimal tasarımı gibi çeşitli uygulama alanlarında ortaya çıkar.

Salkin (1975), Cooper (1981), Floudas (1995), Grossmann (2002), Marchand vd. (2002) ve Wasanapradit vd. (2011) MINLP problemlerini çözmek için oldukça verimli olan klasik hesaplama yöntemlerini önermişlerdir. MINLP problemlerini çözmek için kullanılan deterministik yöntemler

- Genelleştirilmiş bender ayrıştırması
- Dal-Sınır yöntemi
- Dal ve kes algoritması
- Dışsal yaklaşım
- Genelleştirilmiş dışsal yaklaşım
- Eşitlik gevşemeli dışsal yaklaşım
- Eşitlik gevşemeli ve arttırılmış ceza ile dışsal yaklaşım
- Genelleştirilmiş çapraz ayrıştırması
- Fizibilite yaklaşımı

dır. Bu yöntemler, belirli bir problem sınıfına uygulanabilir ve dışbükey olmayan problemlerde global optimuma ulaşabilir. Kirkpatrick vd. (1983), Babu ve Angira (2002), Yan vd. (2004) ve Yiqing vd. (2007) MINLP problemlerinin çözümü için

- Benzetimli Tavlama
- Diferansiyel Evrim Algoritması
- Sıralı Rekabet Algoritması
- Parçacık Sürü Optimizasyonu

yapay zeka optimizasyon yöntemleri kullanmışlardır. Bu sınıftaki algoritmalar genellikle globale yakın optimal çözümler sağlama yeteneğine sahiptir.

MDNLP problemleri ise çoklu yerel optimumlara sahip karmaşık problemlerdir. NLP problemlerine yönelik bilinen yaklaşımların çoğu, karar değişkenlerini sürekli nicelikler olarak ele alır. Fakat, gerçek dünya problemlerine ilişkin matematiksel modeller genellikle kesikli değişkenler içerir. Örneğin, bir dişlinin diş sayısı (tam sayı), bir makinenin seçilip seçilmediği (0-1) ve bir dişlinin standart çap aralığı boyutu (kesikli) verilebilir. MDNLP problemlerini çözmek için birçok yöntem vardır. Fonksiyon yapıları ve değişken tipleri uygun yöntemin seçiminde önemli rol oynar. Literatürde kesikli karar değişkenine sahip problemleri çözmek için uygulanan yöntemler

- Ceza Fonksiyonu Yaklaşımı
- Lagrange Gevşetmesi
- Sıralı Doğrusallaştırma Algoritması
- Yuvarlama Yöntemi

dir. Karar değişkenlerinin her biri için izin verilen kesikli değerlerin numaralandırılması, her zaman kesikli değişken problemlerini çözmektedir. (Arora 2000). Kesikli değişkenler için izin verilen kesikli değerlerin ve kesikli değişkenlerin sayısı arttığında analiz edilecek kombinasyon sayısı da çok hızlı bir şekilde artar. Bu durum problemin hesaplama süresini artırır. Böyle bir durumda çeşitli stratejiler ve sezgisel kurallar kullanarak arama, kombinasyonların olası bir listesine indirgenebilir. Yöntemlerin çoğu sadece doğrusal veya dışbükey problem sınıfı için en uygun çözümü bulur.

MDNLP problemleri için optimizasyon yöntemleri, deterministik ve stokastik arama yöntemleri olarak iki sınıfa ayrılır. Deterministik arama yöntemleri yineleme sürecini temel alır. Dal-Sınır Yöntemi, Ceza Fonksiyonu Yaklaşımı ve Sıralı Doğrusal Programlama Yaklaşımı bu sınıfa aittir. Stokastik arama yöntemleri ise, Rastgele Arama Yöntemi, Karma Yöntemler ve GA'dır. Bu yöntemlerin arama yapmak için probleme özel bilgi gerektirmemesi, problemin doğasıyla ilgili daha az kısıtlama ve daha geniş bir uygulama sağlaması, aramayı tüm arama uzayında yapması nedeniyle global çözümü bulma olasılığı yüksektir. Bu avantajlarının yanı sıra sorunun belirli özelliklerini dikkate almaması, önceki adımlarda elde edilen bilgileri nadiren kullanması, yavaş yakınsama hızları ve düşük doğrulukları nedeniyle arama sürecinin zaman alıcı olması gibi dezavantajları da vardır.

4. OPTİMİZASYON PROBLEMLERİ İÇİN GENETİK ALGORİTMA

Gerçek dünya problemlerinde karar değişkenlerinin tümü sürekli değişkenler olarak ele alınabilir. Bu tür problemler için oluşturulan matematiksel modellerdeki fonksiyonlar türevlenemeyen, doğrusal olmayan vb. yapıda olabilir. Bu yapıdaki problemlerin çözümü için hızlı ve kolay çözüm sağlayan yapay zeka optimizasyon algoritmalarının kullanımı ön plana çıkmaktadır. Yapay zeka optimizasyon algoritmalarından biri olan GA, esnek hesaplama yöntemleri içinde önemli bir yer tutar (Emel ve Taşkın 2002).

4.1 Genetik Algoritma

GA, canlılardaki evrim sürecine ve doğal seçim ilkesine dayanan, genetik süreçleri taklit eden, hem kısıtlı hem de kısıtsız optimizasyon problemlerini çözmek için kullanılan bir esnek hesaplama yöntemidir (Deb 2012). İlk kez John Holland (1975) tarafından önerilen GA, Holland'ın "Adaptation in Natural and Artificial Systems" isimli öncü kitabında yapay zekada uyarlanabilen arama algoritmaları olarak ortaya çıkmıştır. Darwin'in evrim kuramından etkilenen Holland, canlılarda yaşanan genetik süreci bilgisayar ortamında gerçekleştirerek doğal seçim ve genetik evrimden yola çıkıp biyolojik sistemde bireylerin çevreye olan uyumunu örnek alarak araştırmalarında optimali bulmayı amaçlamıştır. 1975 yılında Holland'ın çalışmalarının sonucunda makine öğrenmesi problemlerinde bilgisayar yazılımı modellenerek GA yöntemi ortaya çıkmıştır (Ergül 2010). Bu algoritma ilk olarak De Jong (1975) tarafından optimizasyon problemini çözmek için kullanılmış ve ilk gerçek uygulaması 1984 yılında Schaffer tarafından VEGA'ya uygulanmıştır.

GA, canlılarda yaşanan genetik süreçten esinlenerek oluşturulduğu için genetik biliminde kullanılan bazı kavramlar ve bu kavramların GA'daki karşılıkları bilinmelidir. GA'da kullanılan en küçük yapı birimine *gen* adı verilir. Genler, GA terminolojisinde karar değişkenlerine karşılık gelir. Genlerin alabileceği mümkün değerler *alel* olarak adlandırılır. Genlerin bir araya gelmesiyle *kromozom*, kromozomların bir araya gelmesiyle de *popülasyon* oluşur. Popülasyonda bulunan her bir kromozom *aday çözüm* ve popülasyon ise *aday çözümler kümesine* karşılık gelir (Eğrioğlu vd. 2019).

Çizelge 4.1’de biyoloji ve genetik biliminde kullanılan bazı terimlerin GA terminolojisindeki karşılıkları verilmiştir.

Çizelge 4.1 Biyoloji ve genetik biliminde kullanılan terimlerin GA terminolojisindeki karşılıkları (Cebeci 2021)

Genetik	Optimizasyon
Gen	Karar değişkeni
Alel	Karar değişkeninin tanım kümesi
Birey / Kromozom	Aday çözüm
Popülasyon	Aday çözümler kümesi
Uygunluk fonksiyonu	Amaç fonksiyonu
Uygunluk değeri	Çözüm değeri
Çevre	Kısıtlar
Kuşak / Generasyon / Nesil	Döngü / İterasyon / Yineleme
Evrin süreci	Optimal çözüme doğru döngü
Evrimsel etkenler	Stokastik işlemler
Seçilim, üreme, mutasyon vb.	Seçilim, üreme, mutasyon vb.

GA’nın bilinen optimizasyon yöntemlerinden farklı birçok avantajı vardır. GA, tek bir çözüm yerine her adımda çözümlerden oluşan bir topluluk kullanan popülasyon tabanlı bir arama algoritmasıdır. Çok sayıda en iyi çözümün tek bir adımda bulunabilmesi GA’nın en önemli özelliklerinden biridir. Seçim, çaprazlama ve mutasyon aşamalarında deterministik kuralları yerine olasılıksal geçiş kuralları uygulanır. Parametrelerin kodlanmış değerleriyle hesaplama yapılır. Bu nedenle türevsel veya diğer yardımcı bilgilere ihtiyaç duyulmaz. Arama yaparken doğrudan amaç fonksiyonu değerini kullanır. Popülasyon çeşitliliğinin kaybolması durumunda algoritmanın erken yakınsaması ve yerel optimuma takılması GA’nın dezavantajıdır. Popülasyon çeşitliliğini sağlamak için yüksek uyumlu bireylerin tekrar seçimi önlenmelidir. Popülasyon ne kadar çeşitli ise arama da o kadar global olur. Büyük ölçekli optimizasyon problemlerinde çok sayıda fonksiyon değerlendirmesi yüksek hesaplama maliyeti gerektirmesi de dezavantajdır.

GA tabanlı yöntemlerle kesikli ve/veya sürekli karar değişkenlerinin optimizasyon süreci gerçekleştirilebildiğinden GA çok yönlüdür (Ghasemi vd. 1997). GA’nın her türlü sürekli ve kesikli optimizasyon problemine uygulanabilir olmasından dolayı GA karma optimizasyon problemlerine de uygulanabilir. Hesaplama bakımından basittir ve iyileştirme arayışlarında güçlüdür. Sağlamlık, esneklik ve diferansiyellenebilirlik

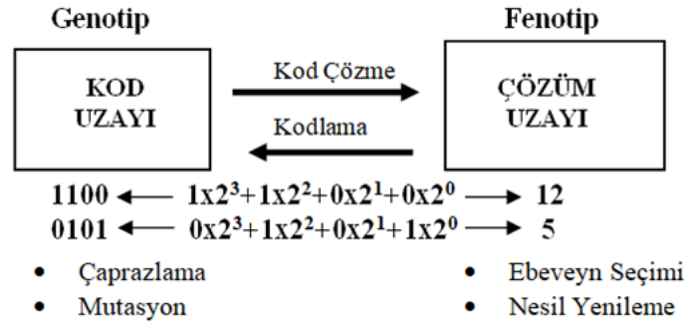
problemi iyi çözüme yeteneği GA'yı kullanışlı kılar. GA ile optimizasyon süreci gerçekleştirilirken optimizasyon probleminin doğrusallık, süreklilik, kesiklilik vb. durumları dikkate alınmadan esnek hesaplama yapılır (Cebeci 2021).

Bütün yapay zeka optimizasyon algoritmasında olduğu gibi ayarlanabilir parametreler GA için de önemlidir. Ayarlanabilir parametreler (tuning parameters) bir modelin performansını optimize etmek için kullanıcı tarafından ayarlanabilen değerlerdir. GA ayarlanabilir parametre değerlerinin probleme uygun biçimde belirlenmiş olması optimizasyon sürecinin başarımını etkiler. Karar değişkenlerinin gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin (seçim, çaprazlama ve mutasyon) uygulanması GA'nın temel çözüm adımları olarak değerlendirilebilir.

4.1.1 Karar değişkenlerinin gösterimi

Bir optimizasyon problemindeki karar değişkenlerinin aldığı değerlere göre GA ile optimizasyonda problemin hangi sayı sisteminde kodlanacağına karar vermek kritik öneme sahiptir. Problemin doğasına ve verilere göre farklı sayı sistemlerinde kodlama yapılabilir. GA'da, karar değişkenlerinin gösterimi farklı kodlamalar kullanılarak gerçekleştirilebilir. Kullanılabilecek başlıca kodlama türleri, ikili (binary), sekizli (Octal), onaltılı (Hexadecimal) ve gri (grey) kodlama ile permütasyon, değer (reel sayı) ve ağaç kodlamasıdır.

Aday çözümlerin gerekli bilgileri içeren bir dizi oluşturması (örneğin 01011 gibi) *kodlama (encoding)* ve kod uzayında oluşturulan dizilerin çözüm uzayında aday çözümlere dönüştürülmesi *kod çözme (decoding)* olarak adlandırılır. Bir başka deyişle kodlama ile fenotipten genotipe, kod çözme ile de genotipten fenotipe dönüşüm yapılır. Kodlamada en önemli ilke kodlamanın tüm olası fenotipleri temsil edebilmesidir. Şekil 4.1'de kodlama ve kod çözme ile kod uzayı ve çözüm uzayında yapılan işlemler görülmektedir.



Şekil 4.1 Kodlama ve kod çözme (Cebeci 2021)

Literatürde en sık kullanılan kodlama türleri ikili, değer kodlama ve permütasyon kodlamasıdır. GA’da, en çok ikili kodlama kullanılmıştır. İkili kodlama, çok boyutlu yüksek hassasiyetli sayısal problemlere uygulandığında, her çözüm için oldukça uzun hesaplama süresi gerektirir (Das ve Pratihar 2018). Sırt çantası problemlerinde (malzeme çantaya konuluyorsa 1 konulmuyorsa 0 gibi) ya da karar değişkenlerinin kesikli olduğu durumlarda tercih edilmektedir. Permütasyon kodlama, permütasyonla çalışılan gezgin satıcı, ulaşım ve çizelgeleme vb. problemlerinin optimizasyonu için yaygın olarak kullanılmaktadır (Satman 2016). Değer kodlaması, tam sayı veya reel sayılar ile ilgilenildiğinde tercih edilmektedir. Michalewicz (1994), değişkenlerin reel ve ikili $\{0,1\}$ sayılarla gösteriminin etkinliğini karşılaştırdığı çalışmalarında, reel sayı gösteriminin daha tercih edilir olduğunu belirtmiştir.

4.1.2 Başlangıç popülasyonunun oluşturulması

Başlangıç popülasyonunun oluşturulması GA’nın ilk adımı olup bu adımda başlangıçta kullanılacak aday çözümler üretilir. Tüm algorithmada başlangıç popülasyonu üretimi yalnızca bir kez kullanılır. GA uygulamalarında genellikle başlangıç popülasyonu rastgele sayı üretici kullanılarak oluşturulur. α , $[0,1]$ aralığından üretilen bir rastgele sayı, lb , değişkenin alt sınırı ve ub , değişkenin üst sınırı olmak üzere, değişkenler için belirlenen $[lb, ub]$ tanım aralığında

$$x = lb + \alpha(ub - lb) \quad (4.1)$$

eşitliğiyle sürekli karar değişkenleri için üretilen çözümlerle başlangıç popülasyonu oluşturulur. Başlangıç popülasyon büyüklüğü (N) ve ikili kodlama ile çalışılması durumunda kromozom uzunluğu (λ) gibi zorunlu ve önemli iki parametre değeri girilmelidir. Rastgele oluşturulan başlangıç popülasyonunda bireylerin uygunluk değerleri hesaplanır. Popülasyondaki bireylerin gelecek nesilde temsil edilip edilmeyeceğine uygunluk fonksiyonu ile hesaplanan uygunluk değerlerine göre karar verilir.

4.1.3 Genetik operatörlerin uygulanması

GA'da, seçim, çaprazlama ve mutasyon olmak üzere üç temel operatör bulunmaktadır. Bu operatörler aşağıda özet biçimde açıklanmıştır.

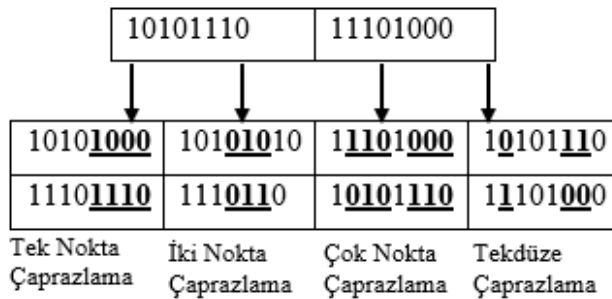
(i) Seçim Operatörü: Seçim operatörünün uygulanmasıyla, popülasyondan uygunluk değeri yüksek olan bireyler seçilerek gelecek nesildeki popülasyonun oluşturulması sağlanır. Seçim sürecinde popülasyon büyüklüğünde eşleşme havuzu oluşturulur. GA'da uygulanan başlıca seçim yöntemleri, Rulet Tekerleği Seçimi, Rank Seçimi, Turnuva Seçimi, Boltzmann Seçimi ve Stokastik Evrensel Örnekleme Seçim yöntemleridir. Sıklıkla kullanılan seçim yöntemleri ise Rulet Tekerleği Seçimi ve Turnuva Seçimi'dir. Rulet Tekerleği Seçiminde, her bireyin bir sonraki popülasyonda yer alma olasılıkları, bireyin uygunluk değerinin toplam uygunluk değerine oranlanarak bulunduğu için literatürde bu seçim yöntemine *orantılı seçim* yöntemi de denilmektedir. Bu seçimde, uygunluk değeri yüksek bireylere daha çok seçilme şansı verilir.

En çok kullanılan bir diğer seçim yöntemi de Turnuva Seçimi'dir. Turnuva seçiminde, popülasyona ait bireylerin uygunluk değerleri karşılaştırılarak seçim yapılır. t , turnuva genişliği olmak üzere, t bireyin mevcut popülasyondan seçimi rastgele yapılır, $t = 2, 3, \dots, N$. $t = 2$ için turnuva seçimi *ikili turnuva seçimi* olarak adlandırılır. Rastgele seçilen t bireyden en iyi olanı yeni popülasyonda yerini alır. Diğer bireyleri seçmek için işlem N kez tekrarlanır. Turnuva genişliği büyük seçilirse, uygunluk değeri yüksek bireylerin seçilme olasılığı yükselir. Popülasyondaki çeşitliliği korumak ve erken

yakınsamayı önlemek için turnuva genişliğinin küçük seçilmesi tercih edilir. Seçim operatörü ile uygunluk değeri yüksek, iyi olarak adlandırılacak çözümler seçilir. Ancak, yeni çözümler üretilemez. Bu nedenle GA, yeni çözümler üretecek operatörlere ihtiyaç duyar. Seçilen çözümlere çaprazlama ve mutasyon işlemleri uygulanarak yeni çözümler elde edilir.

(ii) Çaprazlama Operatörü: Çaprazlama, ebeveynlerden yeni yavru üretme işlemidir. Çaprazlama işleminin amacı, değişkenlerin değerlerini ebeveynler arasında takas yaparak yeni bireyler oluşturmaktır. Bireylerin tümü yerine belli bir yüzdesi arasında çaprazlama gerçekleştirilir. *Çaprazlama olasılığı* (*çaprazlama oranı*) olarak adlandırılan bu değer, P_c ile gösterilir ve genellikle 0.60'dan büyük seçilir.

Uygun bir çaprazlama operatörünün nasıl seçileceği değişken gösterimine bağlıdır. Çaprazlama operatörü kullanılan kodlama türüne göre özelleşir. İkili kodlamanın kullanıldığı problemlerde çaprazlama işlemi tek, iki ya da çok noktadan yapılabilir. Kromozomların belirli bir noktadan kesildiği noktalara *çaprazlama noktası* (*kesme noktası*) adı verilir. Çaprazlama noktası genellikle rastgele olarak belirlenir. Tekdüze çaprazlamada çaprazlama noktalarının sayısı olasılıklı süreçlerle belirlenir. Bu çaprazlama yöntemiyle daha geniş çözüm alanında arama yapılması sağlanır. İkili kodlamada kullanılan çaprazlama yöntemleri şekil 4.2'de görülmektedir. Şekil 4.2'den görüldüğü gibi tek nokta, iki nokta, çok nokta ve tekdüze çaprazlamalar, kromozom üzerindeki genlerin belirli bir parçalanışla yer değiştirmesi biçiminde uygulanır.



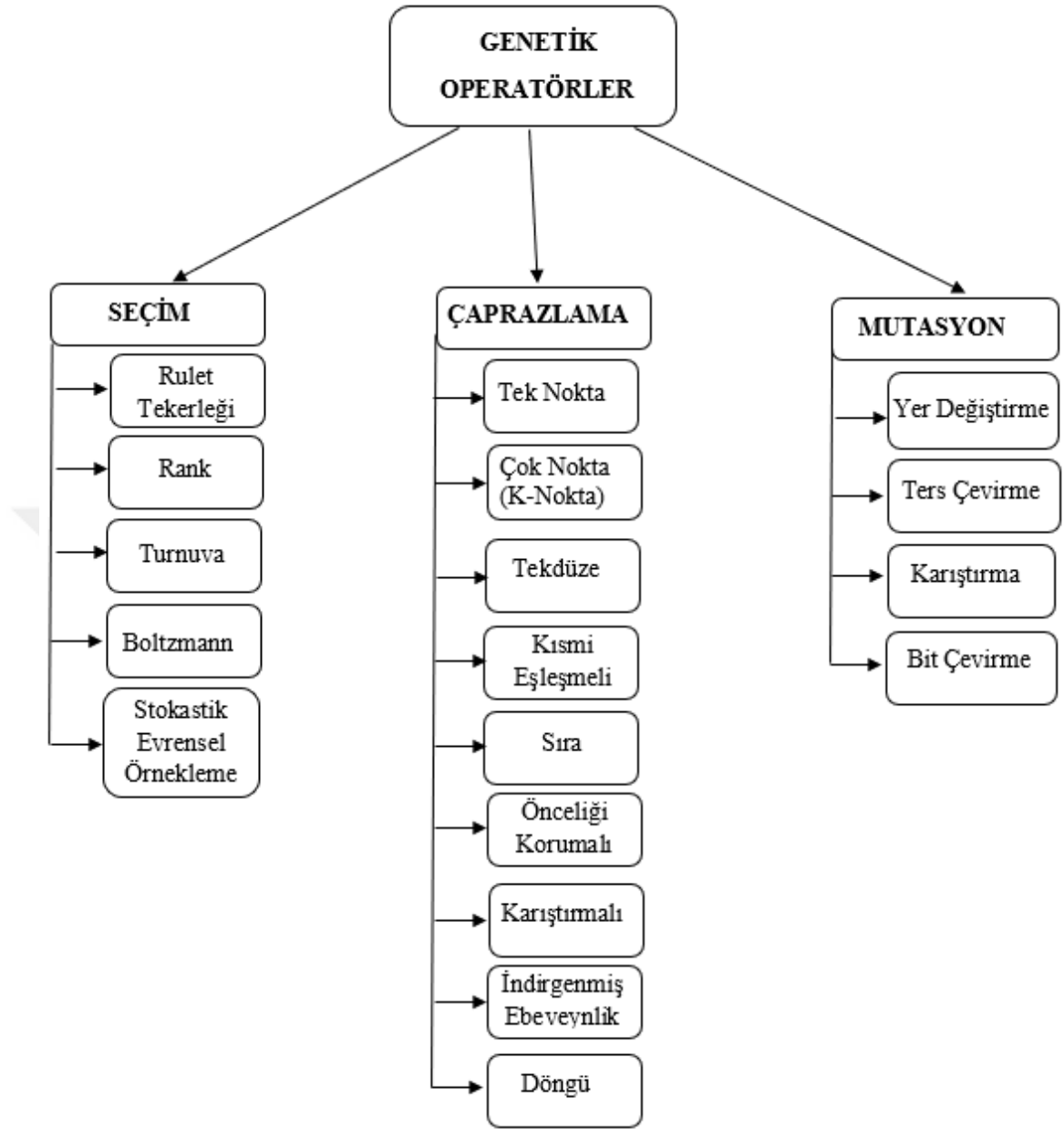
Şekil 4.2 İkili kodlamada kullanılan çaprazlama yöntemleri (Bolat vd. 2004)

Karar deęiřkenlerinin gösteriminde kullanılan bir kodlama türü olan deęer kodlama içeren problemlerde kullanılan başlıca çaprazlama yöntemleri, Düz Çaprazlama, Kesikli Çaprazlama, Simüle Edilmiş İkili Çaprazlama, Doğrusal Çaprazlama, Aritmetik Çaprazlama, Geometrik Çaprazlama, Harmanlanmış Çaprazlama ve Laplace Çaprazlaması'dır (Katoch vd. 2021).

(iii) Mutasyon Operatörü: Mutasyon, kromozomlardaki gen(ler)in belirli bir mutasyon olasılığıyla deęiřtirilmesidir. Bu olasılık deęeri, P_m olarak gösterilir ve genellikle küçük popölasyonlarda %1, büyük popölasyonlarda ise %0.1 oranlarında seçilir. Yüksek mutasyon olasılığı iyi kromozomların yok olma olasılığını arttırır. Mutasyon işlemi popölasyondaki bireylerden çok azına uygulanır. Bu operatör ile kromozomların daha farklı deęerler alması sağlanarak popölasyonda çözüm çeřitlilięi sağlanır. Böylece yeni çözüm bölgelerinde arama yapılır.

İkili kodlanmış GA problemleri için yaygın olarak kullanılan mutasyon yöntemi, Bit Çevirme Mutasyonu'dur. Bit Çevirme Mutasyonunda her bit için [0,1] arasında tekdüze daęılımdan rastgele sayılar çekilir. Çekilen rastgele sayılar ile mutasyon olasılığı karşılaştırılır. Eęer, çekilen rastgele sayı mutasyon olasılıęından küçükse karşılık gelen bit deęeri deęiřtirilir. Aksi halde, bit deęeri aynı kalır.

Kodlama türüne baęlı olarak kullanılan mutasyon yöntemleri farklılık gösterir. Deęer kodlama kullanılan GA problemlerinde uygulanan başlıca mutasyon yöntemleri, Tekdüze Mutasyon, Tekdüze Olmayan Mutasyon, Sınır Deęeri Mutasyonu, Güç Mutasyonu, Yöne Dayalı Üstel Mutasyon ve Polinomsal Mutasyon'dur. Permütasyon kodlama kullanılan GA problemlerinde ise kullanılan başlıca mutasyon yöntemleri, Takas Mutasyonu, Ters Çevirme Mutasyonu, Karıřtırma Mutasyonu ve Yer Deęiřtirme Mutasyonu'dur (Katoch vd. 2021). GA operatörlerinde kullanılan başlıca yöntemlerin řematik olarak gösterimi řekil 4.3'te verilmiřtir.



Şekil 4.3 GA operatörlerinde kullanılan başlıca yöntemler

4.2 Genetik Algoritmanın Teorik Temelleri

GA'nın teorik temelleri, çözümlerin ikili kodlama gösterimine ve kromozomlar arasındaki benzerliklerin keşfedilmesine olanak tanıyan bir şablon olan şema kavramına dayanmaktadır (Michalewicz 1994). Holland (1975), GA'nın teorik analizini göstermek için şema teoremini kullanmıştır. Holland'ın şema teoremi, GA'nın başarısını açıklamak için kullanılan en başarılı teoremdir (İşçi ve Korukoğlu 2003). Bir şema, genlerin

alfabesine “*” sembolü eklenerek oluşturulur. İkili kodlamada şema gösterimi için {0, 1, *} alfabeti kullanılır. Burada “*” sembolü *sıfır veya bir ile değiştirilebilir* anlamına gelir. Holland, şema teoremi ile genetik operatörler uygulanarak mevcut popülasyonda temsil edilen her şemanın eleman sayısının diğer şemalardan bağımsız olarak nispi uygunluk değeriyle orantılı olarak azalacağını veya artacağını göstermeyi amaçlamıştır (Karaboğa 2020). Matematiksel terimler kullanarak şema için *tanım uzunluğu* ve *şemanın derecesi* notasyonları ile şema teoremini ispat etmiştir (Michalewicz 1994).

Örneğin, uzunluğu 10 bit olan kromozomlar, (*1*1100100) şemasının bir örneğidir. Bu şema {(0101100100), (0111100100), (1101100100), (1111100100)} olmak üzere dört kromozomla eşleşir. (0101100100) şeması yalnızca bir kromozomu temsil ederken (******) şeması 10 bit uzunluğundaki tüm kromozomları temsil eder. Genel olarak, r şema şablonundaki * sembolünün sayısı olmak üzere, her şema 2^r kromozomla eşleşir. Diğer taraftan, m uzunluğundaki her bir kromozom, 2^m tane şemayla eşleştirilir. Örneğin, kromozom uzunluğu 10 olan (1001110001) kromozomu

(1001110001)
 (*001110001)
 (1*01110001)
 ⋮
 (100111000*)
 (**01110001)
 (*0*1110001)
 ⋮
 (10011100**)
 (***1110001)
 ⋮
 (******)

olmak üzere 2^{10} tane şema ile eşleştirilir. N büyüklüğünde bir popülasyon, 2^m ile $N.2^m$ arasında farklı şema ile temsil edilebilir. Farklı şemaların farklı özellikleri vardır. Şemanın iki önemli özelliği; (i) sabit ilk ve son pozisyonları arasındaki uzunluk, ve (ii) sabit olarak tanımlayıcı konumların sayısıdır.

Bir S şemasının derecesi, şemada mevcut olan 0 ve 1 konumlarının sayısıdır. Başka bir deyişle, şemadaki sabit pozisyonların sayısıdır ve $o(S)$ ile gösterilir. Örneğin, her birinin uzunluğu 10 olan S_1 , S_2 ve S_3 şemalarının derecesi

$$\begin{aligned} S_1 &= (**001*110) \\ S_2 &= (****00**0*) \\ S_3 &= (11101**001) \end{aligned}$$

$o(S_1) = 10 - 4 = 6$, $o(S_2) = 10 - 7 = 3$ ve $o(S_3) = 10 - 2 = 8$ 'dir. S_3 şeması en belirgin olanıdır. S şemasının tanımlayıcı uzunluğu, şemanın sabit ilk ve son konumları arasındaki uzaklıktır ve $\delta(S)$ ile gösterilir. Üç şema için tanımlayıcı uzunluklar, $\delta(S_1) = 10 - 4 = 6$, $\delta(S_2) = 9 - 5 = 4$ ve $\delta(S_3) = 10 - 1 = 9$ 'dur.

Bir popülasyonda, iyi çözümleri gösteren bir şemaya uyan kromozomların sayısının bir sonraki nesilde artması istenir. Bir nesilden diğer nesile geçişte popülasyonda yer alan şemaların beklenen sayılarını seçim, çaprazlama ve mutasyon operatörleri belirler (Taştan 2012). Şemanın bir diğer özelliği, t . yinelemedeki uygunluk değeridir, $eval(S, t)$. Popülasyonda, t . yinelemede bir S şeması tarafından eşleştirilen p tane $\{v_{i1}, v_{i2}, \dots, v_{ip}\}$ kromozom olduğunu varsayalım. S şemasına uyan bu kromozomların ortalama uygunluk değerine S şemasının uygunluğu denir ve şema uygunluğu

$$eval(S, t) = \sum_{j=1}^p eval(v_{ij}) / p, \quad i = 1, 2, \dots, N \quad (4.2)$$

biçiminde hesaplanır. Seçim işleminden sonra her bir kromozom uygunluk değerine göre bir veya daha fazla kez kopyalanır ya da kopyalanmaz. t . yinelemede tüm popülasyonun

uygunluk değerleri toplamı $F(t) = \sum_{i=1}^N eval(v_i)$ olmak üzere v_i kromozomu

$$p_i = eval(v_i) / F(t), \quad i = 1, 2, \dots, N \quad (4.3)$$

olasılığı ile seçilir. Seçim işleminden sonra $(t+1)$. yinelemede popülasyonda S şemasına uyan

$$\zeta(S, t+1) = \zeta(S, t) \times N \times eval(S, t) / F(t) \quad (4.4)$$

kromozomun bulunması beklenir. Burada, N , popülasyon büyüklüğü olmak üzere, popülasyonun ortalama uygunluk değeri $F(t)$, $\overline{F(t)} = F(t)/N$ olarak alınırsa eşitlik (4.4)

$$\zeta(S, t+1) = \zeta(S, t) eval(S, t) / \overline{F(t)} \quad (4.5)$$

biçiminde yazılır. Başka bir deyişle, şemanın uygunluğunun popülasyonun ortalama uygunluğuna oranı kadar popülasyondaki kromozomların sayısı artar. Bir şemaya uyan kromozomların sayısı, ilgili şema ortalamanın üstünde uygunluk değerine sahip ise sonraki nesilde artar, ortalamanın altında uygunluk değerine sahip ise azalır ve ortalama bir uygunluk değerine sahip ise aynı düzeyde kalır. Eğer bir S şemasının ortalamanın $\% \varepsilon$ kadar üstünde bir uygunluk değerine sahip olduğu varsayılırsa

$$\varepsilon = \frac{eval(S, t) - \overline{F(t)}}{\overline{F(t)}} \quad (4.6)$$

olmak üzere

$$\zeta(S, t) = \zeta(S, 0)(1 + \varepsilon)^t \quad (4.7)$$

dir. Ortalamanın üstünde uygunluk değerine sahip şema için $\varepsilon > 0$ ve ortalamanın altında uygunluk değerine sahip şema için $\varepsilon < 0$ olur. Bu geometrik ilerleme denklemi ile böyle bir şemanın sonraki nesillerde katlanarak artan sayıda kromozom aldığı söylenebilir.

Seçim işlemi, tek başına arama alanında dikkate alınacak herhangi bir yeni nokta (potansiyel çözüm) getirmez, bir ara popülasyon oluşturmak için yalnızca bazı kromozomları kopyalar. Popülasyona yeni bireylerin dahil edilmesi çaprazlama ve mutasyon operatörleri ile sağlanır. Çaprazlama operatörünün popülasyondaki beklenen şema sayısı üzerindeki etkisi bir şemanın tanımlayıcı uzunluğuna ve çaprazlama olasılığına bağlıdır. Bir şemanın tanımlayıcı uzunluğu o şemanın yok olması ve hayatta

kalması olasılığında önemli bir rol oynar. Genel olarak, çaprazlama konumu, $m-1$ olası konumlar arasından eşit biçimde seçilir. S şemasının yok olma ve hayatta kalma olasılığı sırasıyla

$$P_d(S) = \frac{\delta(S)}{m-1} \quad (4.8)$$

ve

$$P_s(S) = 1 - \frac{\delta(S)}{m-1} \quad (4.9)$$

biçiminde verilir. P_c çaprazlama olasılığı olmak üzere, sadece bazı kromozomların çaprazlandığı göz önüne alınırsa bir şemanın hayatta kalma olasılığı

$$P_s(S) = 1 - P_c \frac{\delta(S)}{m-1} \quad (4.10)$$

olur. Seçim ve çaprazlama operatörlerinin etkileri birleştirildiğinde üreme şeması büyüme denkleminin yeni hali

$$\zeta(S, t+1) \geq \zeta(S, t) \text{eval}(S, t) / \overline{F(t)} \left[1 - P_c \frac{\delta(S)}{m-1} \right] \quad (4.11)$$

dır. Bir sonraki nesilde, bir S şemasıyla eşleşen kromozomların beklenen sayısı, şemayla eşleşen gerçek kromozom sayısının ve şemanın göreceli uygunluk değeri ile tanımlayıcı uzunluğunun bir fonksiyonudur. Kısa tanımlama uzunluğuna sahip ortalamanın üstünde şemalar artan oranlarda örneklenir.

Mutasyon operatörü, kromozom içindeki biti P_m olasılığıyla sıfırdan bire veya birden sıfıra rastgele değiştirir. Tek bir bitin değişme olasılığı P_m olduğundan, tek bir bitin hayatta kalma olasılığı da $1 - P_m$ 'dir. Tek bir mutasyon diğer mutasyonlardan bağımsızdır. Dolayısıyla, bir S şemasının bir mutasyondan sonra hayatta kalma olasılığı

$$P_S(S) = (1 - P_m)^{o(S)} \quad (4.12)$$

dır.

Küçük mutasyon olasılıkları $P_m \ll 1$ için bu olasılık yaklaşık olarak

$$P_s(S) \approx 1 - o(S)P_m \quad (4.13)$$

olarak kabul edilir. Seçim, çaprazlama ve mutasyon operatörlerinin etkileri birleştirildiğinde üreme şeması büyüme denkleminin son hali

$$\xi(S, t+1) \geq \xi(S, t) \text{eval}(S, t) / \overline{F(t)} \left[1 - P_c \frac{\delta(S)}{m-1} - o(S)P_m \right] \quad (4.14)$$

olur. Büyüme denkleminin sonucu, kısa, düşük dereceli, ortalamanın üstündeki şemalar, GA'nın sonraki nesillerinde üstel olarak artan oranda örnek elde ederler. Bu teoremin doğrudan bir sonucu, GA'nın arama uzayını kısa, düşük dereceli şemalarla keşfettiği ve daha sonra çaprazlama sırasında bilgi alışverişi için kullandığıdır.

GA, yapı blokları olarak adlandırılan kısa, düşük dereceli, yüksek performanslı şemaların bir araya getirilmesiyle optimuma yakın performans arar. Şema teoremi üzerine kurulan yapı blokları hipotezi, GA'nın başarısındaki temel hipotezdir. Bu hipotez, GA'nın kodlama probleminin performansı bakımından kritik olduğunu ve böyle bir kodlamanın kısa yapı blokları fikrini karşılaması gerektiğini ileri sürmektedir (Michalewicz 1994).

4.3 Tek Amaçlı Optimizasyon Problemleri için Genetik Algoritma

GA'da optimize edilecek amaç sayısı tek olduğunda, tek amaçlı GA olarak adlandırılır. Tek amaçlı GA'nın işlem süreci çok amaçlı GA'dan farklı değildir. Tek amaçlı optimizasyon problemlerinin çözümü için uygulanan GA yönteminin algoritmik adımları aşağıdaki gibidir.

Adım 0 : t , yineleme sayısı olmak üzere, GA'nın ayarlanabilir parametreleri, popülasyon büyüklüğü (N), kromozom uzunluğu (λ), çaprazlama olasılığı (P_c), çaprazlama indeksi (η_c), mutasyon olasılığı (P_m), mutasyon indeksi (η_m) ve maksimum yineleme sayısı (n_{gen}), uzman görüşüne göre belirlenir, $t = 1, 2, \dots, n_{gen}$.

Adım 1: Başlangıç popülasyonu oluşturulur ve uygunluk fonksiyonu değerleri hesaplanır.

Adım 2: Bireylerin uygunluk fonksiyonu değerlerine göre ebeveynler seçilir ve eşleşme havuzu oluşturulur.

Adım 3: Eşleşme havuzundaki ebeveynler çaprazlanarak yavru bireyler üretilir.

Adım 4: Çaprazlama sonucu üretilen bazı yavru bireylere mutasyon uygulanır.

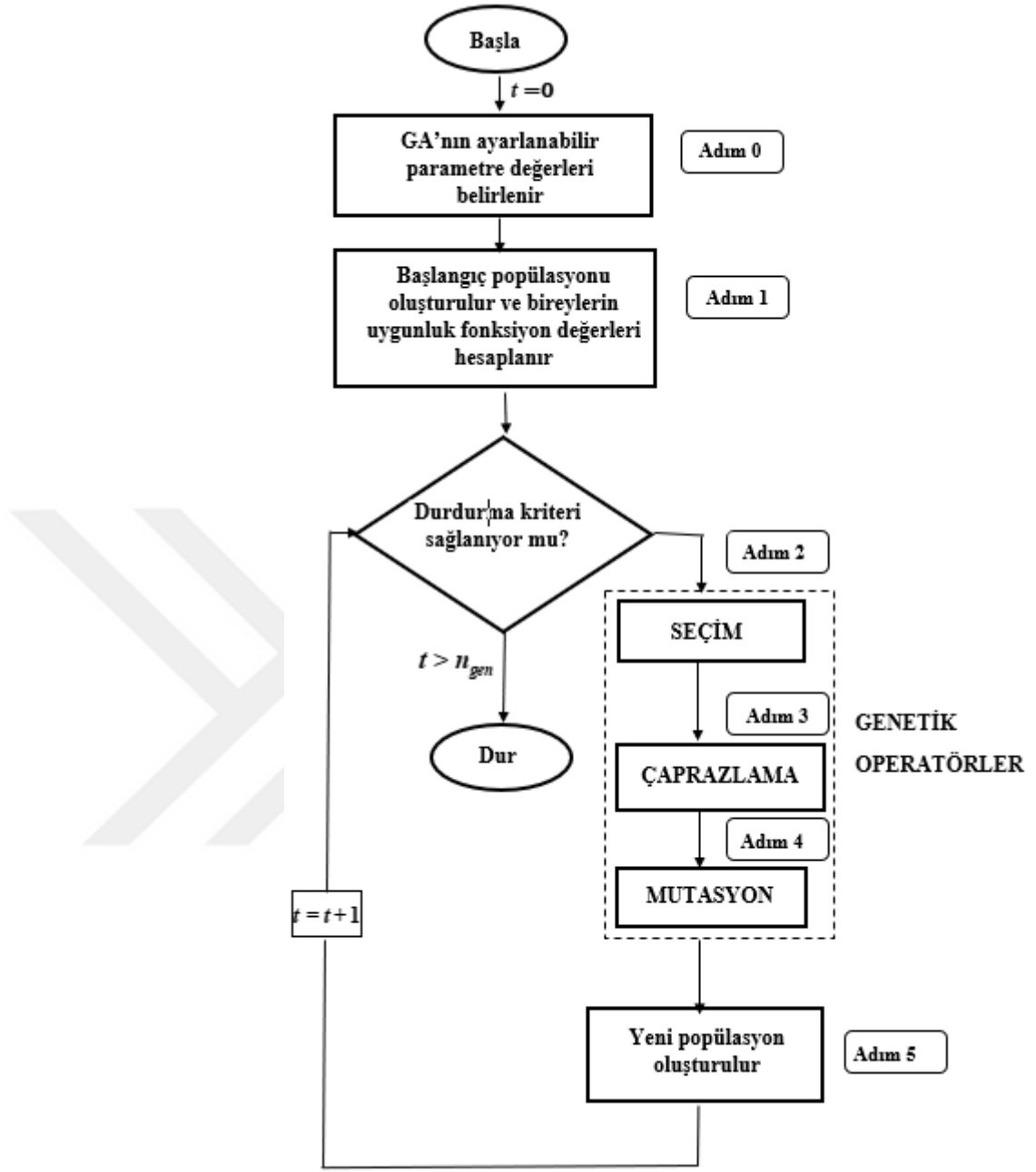
Adım 5: Yeni popülasyonu oluşturacak olan bireyler belirlenir.

Adım 6: Başlangıçta belirlenen maksimum yineleme sayısına ulaşılmış ise algoritma sonlandırılır. Aksi halde, *Adım 2*'ye gidilir.

Ebeveyn ve yavru popülasyonlardan oluşan birleştirilmiş popülasyonda ($2N$), popülasyon büyüklüğünün sabit tutulması için mevcut popülasyondaki bazı ebeveynlerin atılması ve yerlerine yavruların konulması gerekir. Bu işlem, *popülasyon yenileme* olarak adlandırılır. Popülasyon yenilenirken uygunluk değeri yüksek olan bireyler seçilir ve popülasyondaki çeşitliliğin de sürdürülmesi istenir.

GA'nın işlem süreci bazı durdurma kriterlerine göre sonlandırılır. Bu kriterler, algoritmanın işlem sürecinde önceden tanımlanmış global optimuma ulaşılması, GA'nın başlangıcında tanımlanan maksimum yineleme sayısına ulaşılması ve uygunluk değerinin durağanlaşıp sürekli olarak aynı değere yakınsaması olarak tanımlanabilir. Uygulamalarda genellikle durdurma kriteri olarak maksimum yineleme sayısı kullanılır. Kullanıcı tarafından belirlenen bu değere ulaşıldığında algoritma durur. Bu ayarlanabilir parametrenin varsayılan değeri genellikle 100 olarak tanımlanır, $n_{gen} = 100$.

GA'nın algoritmik adımları ve işlem basamakları şekil 4.4'te verilen GA akış şemasında görülmektedir.



Şekil 4.4 GA'nın akış şeması

Ayarlanabilir parametre değerlerinin uygun biçimde belirlenmiş olması GA'nın performansı üzerinde oldukça etkilidir. Bu parametrelerin farklı kombinasyonları farklı optimal çözümlerle sonuçlanabilir. Kabul edilebilir çözümler elde etmek için ayarlanabilir parametreleri doğru biçimde tanımlamak zahmetli bir iştir. De Jong (1975), Grefenstette (1986) ve Schaffer vd. (1989) tarafından yürütülen çalışmalarda, GA

ayarlanabilir parametrelerinin seçiminde kılavuz olarak kullanılabilecek bilgilere ulaşmak mümkündür.

Çizelge 4.2 GA için literatürde önerilen ayarlanabilir parametre değerleri

Ayarlanabilir Parametreler			
Literatür	N	P_c	P_m
<i>De Jong (1975)</i>	50-100	0.60	0.001
<i>Grefenstette (1986)</i>	30	0.95	0.01
<i>Schaffer (1989)</i>	20-30	0.75-0.95	0.005-0.01

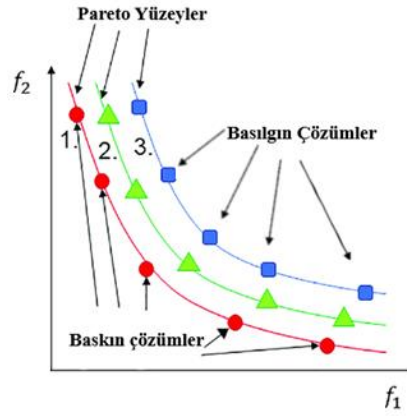
Çizelge 4.2’de, GA için önerilen ayarlanabilir parametre değerleri verilmiştir. Popülasyon büyüklüğü (N) çok küçük seçildiğinde GA’nın performansı düşer, çok büyük seçildiğinde ise algoritmanın çalışma süresi uzar. Çaprazlama olasılığı (P_c) ve mutasyon olasılığı (P_m) seçimi de çok önemlidir. Çok küçük mutasyon olasılığı nedeniyle algoritma yerel optimuma takılabilir ya da çok büyük mutasyon olasılığı seçilmesi durumunda algoritma optimuma yakınsamayabilir. Nesil sayısının ya da maksimum yineleme sayısının belirlenmesi zordur. Küçük yineleme sayısı optimum çözüme ulaşmak için genetik ilerlemeye engel olurken büyük yineleme sayısı hesaplama süresinin uzamasına yol açar. Bu nedenle, ayarlanabilir parametreler probleme uygun biçimde seçilmelidir.

4.4 Çok Amaçlı Optimizasyon Problemleri için Genetik Algoritma

MOGA, MOO problemlerini çözmek için geliştirilmiş EA sınıfında yer almaktadır. Birbiriyle çelişen amaç fonksiyonunu optimize eden Pareto çözümleri bulmakla ilgilenir. MOGA’da farklı amaçlardan hesaplanan ve birbirine karşı ödün vermek zorunda olan bir dizi çözüm vardır. Bu çözümlere, *Pareto çözümler* denir. Pareto çözümlerden hangisinin daha iyi olduğu, karar vericinin kriterlere verdiği önceliğe göre belirlenir. Örneğin, bir fabrikada ürün imalatında harcamalar minimize edilmek istenirken, kar maksimize edilmek istenir. Harcamalar düşük tutulmaya çalışılırken, karın yüksek olmasının istenmesi birbiriyle çelişen amaçlar arasında bir ödünleşme sağlayacak biçimde üretim yapılmaya çalışılır.

Klasik optimizasyon yöntemleri, MOO problemini tek amaçlı bir optimizasyon problemine dönüştürmeyi önerir. EA'lar çözüm popülasyonu ile çalıştığından, Pareto yüzeye bölgeye doğru ilerlemeye vurgu yapan EA'lar ile tek bir hesaplamada Pareto çözüm kümesi elde edilebilir.

Şekil 4.5'te, iki amaçlı minimizasyon problemi için Pareto yüzeyi ile baskın ve basılgin çözüm kümeleri görülmektedir. Çok amaçlı optimizasyonda amaç, şekil 4.5'te "1." ile ifade edilen *Rank 1* baskın çözüm kümesinin çözüm çeşitliliği sağlanacak biçimde elde edilmesidir. Bu süreçte, Pareto tabanlı esnek hesaplama yöntemleri ile optimizasyon tercih edilir.



Şekil 4.5 Pareto yüzeyi, basılgin ve baskın çözümler kümesi

Pareto tabanlı yöntemler sınıfında yer alan MOGA'da, NSGA yöntemine dayanan NSGA-II ve NSGA-III yöntemleri kullanılır (Cebeci 2021). Bu tez çalışmasında, çok amaçlı optimizasyon problemlerinin çözümü için GA'ya dayalı Pareto çözüm elde edilmesini sağlayan esnek hesaplama yöntemlerinden olan NSGA-II ve NSGA-III yöntemlerinin uygulanması tercih edilmiştir.

4.4.1 NSGA-II yöntemi

İlk olarak Goldberg (1989) tarafından önerilen NSGA-II, Srinivas ve Deb'in (1994) NSGA uygulamaları ile hayata geçmiştir. Srinivas ve Deb (1994) tarafından önerilen NSGA, Pareto çözüm kümesinin elde edildiği ilk EA'lardan birisidir. Fakat NSGA

yöntemiyle çok amaçlı optimizasyonda baskın sıralamanın hesaplama zorluğu, seçkinlik eksikliği ve çözüm çeşitliliğinin yeterince sağlanamaması gibi durumlar söz konusudur. NSGA'nın geliştirilmiş versiyonu olan NSGA-II ile tüm bu sorunların üstesinden gelinerek çözüm kümesinde çeşitliliği ve Pareto yüzeye yakınsamanın sağlandığı görülmüştür (Deb vd. 2002).

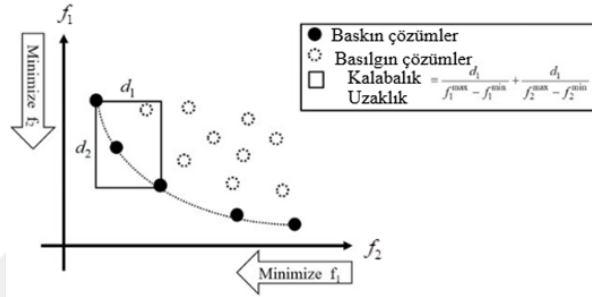
NSGA-II, Deb vd. (2002) tarafından geliştirilen bir sezgisel MOGA'dır. NSGA-II ile klasik MOGA yöntemleri kıyaslandığında, NSGA-II'nin daha çok çözüm çeşitliliğiyle Pareto yüzeye daha iyi bir yakınsama sağladığı görülmüştür (Deb vd. 2002). Arama, bir başlangıç popülasyonu ile başlar ve algoritma optimum çözümleri bulmak için genetik operatörleri kullanarak arama uzayında olasılıksal popülasyon tabanlı bir keşif uygular. Ardışık nesillerdeki bireysel çözümlerin uygunluğu, seçim, çaprazlama ve mutasyon operatörlerinin uygulanmasıyla artar (Golberg 1989). Süreç, önceden tanımlanmış sonlandırma koşulu sağlandığında durur. Pareto çözümü bulmak için gereken hesaplama süresi, karar değişkenleri ve amaç fonksiyonu sayısı arttıkça artar.

NSGA-II'nin diğer MOGA'lara göre üstünlüğü, *hızlı baskın sıralama* (*fast nondominated sorting*) ve *kalabalık uzaklığı* (*crowding distance*) yaklaşımlarıdır. Bu yaklaşımlar, baskın yüzeyin oluşturulmasında kullanılıp Pareto çözüm kümesinde farklı çözümler elde edilmesini sağlar. Böylece Pareto yüzeye daha çabuk ulaşılabilir (Türkşen 2011).

Hızlı baskın sıralamada, ilk baskın yüzeydeki çözümleri belirlemek için N büyüklüğündeki bir popülasyonda her çözüm, popülasyondaki diğer tüm çözümler ile karşılaştırılır. İlk baskın yüzeydeki tüm bireyler bulunduktan sonra bir sonraki yüzeydeki bireyleri bulmak için ilk yüzey çözümleri geçici olarak göz ardı edilir ve aynı işlemler tekrarlanır. Bireyler baskınlık derecelerine göre farklı yüzeylerde yer alır. Hızlı baskın sıralama işleminin algoritmik adımları Deb vd.'nin (2002) çalışmalarında görülebilir.

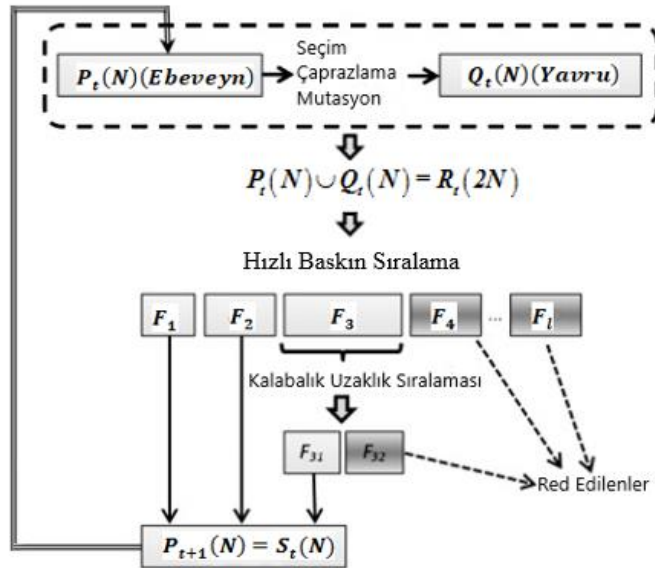
Pareto çözümlerin sıralanmasında, her bir çözüm için baskınlık sayacıyla ilgili çözümün baskın olduğu çözüm kümesi kullanılır. Çözümlerin sıralanması NSGA'ya göre daha hızlı olduğu için tanımlanan sıralama algoritması *hızlı baskın sıralama algoritması* olarak adlandırılır.

Kalabalık uzaklığı, popülasyondaki belirli bir i , $i = 1, 2, \dots, N$, çözümünü çevreleyen çözümlerin yoğunluğunun tahmini için her bir amaç fonksiyonu için i çözümünün her iki tarafındaki iki çözümün ortalama öklid uzaklığı biçiminde hesaplanır. Bu uzaklık değeri, en yakın komşular kullanılarak oluşturulan küboidin çevre uzunluğunun tahminiyle elde edilir (Deb 2001). Kalabalık uzaklık, aynı Pareto yüzeydeki baskın çözümler için hesaplanır. Şekil 4.6’da aynı yüzeyde yer alan baskın çözümler için kalabalık uzaklık hesaplaması gösterilmiştir.



Şekil 4.6 Baskın sıralama ve kalabalık uzaklık yaklaşımları (Choi ve Kim 2019)

Pareto çözüm kümesinde çözümlerin yayılımı ve çeşitliliği kalabalık uzaklık yaklaşımı ile sağlanır. Şekil 4.7’de NSGA-II’nin hızlı baskın sıralama ve kalabalık uzaklık hesaplama işleyişi görülmektedir.



Şekil 4.7 NSGA-II’nin hızlı baskın sıralama ve kalabalık uzaklık hesaplama işleyişi (El-Kribi vd. 2013)

Baskınlık kriterine göre sıralanan bireylere kalabalık uzaklık değeri atandıktan sonra ebeveynler, ikili turnuva seçimi kullanılarak seçilir. Seçilen ebeveynlerden *Simüle Edilmiş İkili Çaprazlama (Simulated Binary Crossover-SBX)* operatörü kullanılarak yavrular üretilir ve üretilen bazı yavrulara *Polinomsal Mutasyon operatörü* uygulanır. NSGA-II yönteminde kullanılan çaprazlama ve mutasyon operatörleri aşağıda tanımlanmıştır.

(i) Simüle Edilmiş İkili Çaprazlama (Simulated Binary Crossover-SBX): SBX operatörü sürekli değer alan karar değişkenleri için uygulanır. SBX operatörü altında iki ebeveyn (y_1 ve y_2) çözümden yavru (c_1 ve c_2) çözümleri hesaplamak için algoritma adımları aşağıdaki gibidir (Deb ve Agrawal 1999).

Adım 1: $[0,1]$ aralığında rastgele bir sayı oluşturulur, u .

Adım 2: Polinomsal olasılık dağılımı kullanılarak bir β_q parametresi bulunur.

$$\beta_q = \begin{cases} (u\alpha)^{\frac{1}{\eta_c+1}} & , \quad u \leq \frac{1}{\alpha} \\ \left(\frac{1}{2-u\alpha}\right)^{\frac{1}{\eta_c+1}} & , \quad d.y. \end{cases} \quad (4.15)$$

Burada $\alpha = 2 - \beta^{-(\eta_c+1)}$ ve $\beta = 1 + \frac{2}{y_2 - y_1} \min[(y_1 - y_{lb}), (y_{ub} - y_2)]$ 'dir. y çözümünün

$[y_{lb}, y_{ub}]$ aralığında değiştiği varsayılır. η_c çaprazlama parametresi, SBX için dağılım indeksidir ve negatif olmayan herhangi bir değer alabilir. Küçük bir η_c değeri, ebeveynlerden uzak çözümlerin yavru çözümler olarak oluşturulmasına izin verirken, büyük bir η_c değeri de ebeveyne yakın çözümlerin yavru çözümler olarak oluşturulmasına olanak sağlar.

Adım 3: Yavru çözümler

$$\begin{aligned} c_1 &= 0.5[(y_1 + y_2) - \beta_q |y_2 - y_1|] \\ c_2 &= 0.5[(y_1 + y_2) + \beta_q |y_2 - y_1|] \end{aligned} \quad (4.16)$$

biçiminde hesaplanır. Burada, $y_1 < y_2$ olduğu varsayılmaktadır. Birden fazla değişkeni ele almak için her değişken eşit olasılıkla seçilir.

(ii) Polinomsal Mutasyon Operatörü: Deb ve Agrawal (1999) tarafından önerilen Polinomsal Mutasyon operatöründe, kullanıcı tarafından mutasyon indeks parametresi η_m tanımlanır. Bu mutasyon operatörü, bir ebeveyn çözüm y 'nin yakınında, bir yavru çözüm c 'yi oluşturmak için polinom olasılık dağılımını kullanır. Bir $y \in [y_{lb}, y_{ub}]$ çözümü için aşağıdaki adımlar uygulanır.

Adım 1: $[0, 1]$ aralığında rastgele bir sayı oluşturulur, u .

Adım 2: δ_q parametresi

$$\delta_q = \begin{cases} \left[2u + (1-2u)(1-\delta)^{\eta_m+1} \right]^{\frac{1}{\eta_m+1}} - 1 & , \quad u \leq 0.5 \\ 1 - \left[2(1-u) + 2(u-0.5)(1-\delta)^{\eta_m+1} \right]^{\frac{1}{\eta_m+1}} & , \quad d.y. \end{cases} \quad (4.17)$$

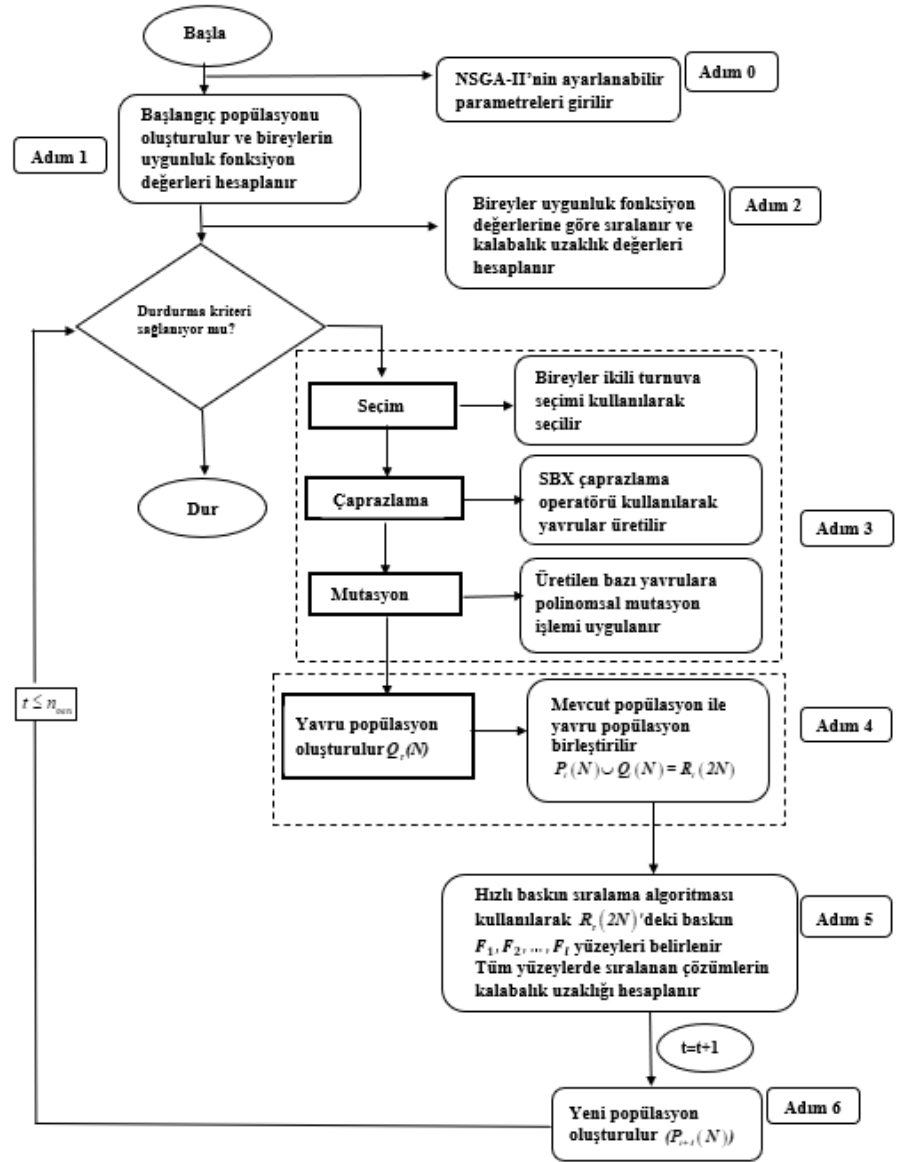
biçiminde hesaplanır. Burada, $\delta = \min[(y - y_{lb}), (y_{ub} - y)] / (y_{ub} - y_{lb})$ 'dir. η_m mutasyon parametresi, mutasyonun dağılım indeksi olup negatif olmayan bir değer alır.

Adım 3: Mutasyona uğrayan yavru

$$c = y + \delta_q (y_{ub} - y_{lb}) \quad (4.18)$$

eşitliği ile elde edilir.

Akış şeması şekil 4.8'de verilen NSGA-II yönteminin algoritmik adımları aşağıda tanımlanmıştır.



Şekil 4.8 NSGA-II'nin akış şeması

NSGA-II Algoritması:

Adım 0: t , yineleme sayısı olmak üzere, NSGA-II'nin ayarlanabilir parametreleri, popülasyon büyüklüğü (N), girdi değişken sayısı (karar değişkeni) (p), çaprazlama olasılığı (P_c), çaprazlama indeksi (η_c), mutasyon olasılığı (P_m), mutasyon indeksi (η_m) ve maksimum yineleme sayısı (n_{gen}), uzman görüşüne göre belirlenir, $t = 1, 2, \dots, n_{gen}$. Karar değişkenleri için belirlenen tanım aralıklarından karar değişkenleri değerleri tanımlanır.

Adım 1: Rastgele oluşturulan başlangıç popülasyonunda bireylerin uygunluk fonksiyonu değerleri hesaplanır.

Adım 2: Bireyler uygunluk fonksiyon değerlerine göre sıralanır ve kalabalık uzaklık değerleri hesaplanır. Bireyler sıralamaya ve kalabalık uzaklığa göre seçildiğinden popülasyondaki tüm bireylere bir kalabalık uzaklık değeri atanır.

Adım 3: Genetik operatör uygulamaları seçim, çaprazlama ve mutasyon aşamalarından oluşur. *Adım 0*'da belirtilen ayarlanabilir parametre değerleri kullanılarak genetik operatör işlemleri aşağıdaki gibi uygulanır.

- Baskınlık kriterine göre sıralanan bireylere kalabalık uzaklık değeri atandıktan sonra seçim, bir kalabalık karşılaştırma operatörü ($\prec$) kullanılarak gerçekleştirilir. Bireyler, ikili turnuva seçimi kullanılarak seçilir. Seçim, iki kritere göre yapılır. İlki ve en önemlisi, çözümlerin bulunduğu yüzey veya sıradır. Daha düşük sıralamaya sahip bireyler seçilir. İkinci olarak, iki bireyin sıralaması aynı ise, kalabalık uzaklığı karşılaştırılır. Kalabalık uzaklığı daha büyük olan bireyler seçilir.
- Seçilen ebeveynlerden SBX operatörü kullanılarak yavrular üretilir.
- SBX operatörü ile üretilen bazı yavrulara Polinomsal mutasyon operatörü uygulanır.

Adım 4: $t = 1, 2, \dots, n_{\text{gen}}$ olmak üzere, N ve $2N$ bireyli yeni popülasyon oluşturulur.

- Yeni $Q_t(N)$ yavru popülasyonu oluşturulur.
- Mevcut popülasyon ile oluşturulan yavru popülasyon birleştirilir, $P_t(N) \cup Q_t(N) = R_t(2N)$. Oluşturulan $R_t(2N)$ popülasyonundaki tüm çözümlerin birlikte değerlendirilmesi ile seçkinlik işlemi gerçekleştirilir.

Adım 5: Hızlı baskın sıralama algoritması kullanılarak $R_t(2N)$ 'deki baskın $F_1, F_2, \dots, F_l$ yüzeyleri belirlenir, $t = 1, 2, \dots, n_{\text{gen}}$. Tüm yüzeylerde sıralanan çözümlerin kalabalık uzaklığı hesaplanır. Kalabalık uzaklık değerlerine göre başlangıçta belirlenen popülasyon sayısı (N) elde edilecek biçimde sıralı yüzeylerden bireyler seçilir.

Adım 6: İkili Turnuva Seçim Yöntemi kullanılarak $R_t(2N)$ ’deki ilk N birey seçilerek yeni nesil ebeveyn popülasyon $P_{t+1}(N)$ oluşturulur, $t = 1, 2, \dots, n_{\text{gen}}$. Başlangıçta belirlenen yineleme sayısına ulaşılmış ise algoritma sonlandırılır. Aksi halde *Adım 3*’e gidilir.

4.4.2 NSGA-III yöntemi

Baskın Sıralı Genetik Algoritma-III (Nondominated Sorting Genetic Algorithm-III / NSGA-III), referans noktalarına dayalı yeni bir çeşitlilik koruma yöntemine ve NSGA-II’nin baskın sıralama yaklaşımına dayanır. İlk olarak Deb ve Jain (2014) tarafından sunulan Pareto tabanlı algoritmalarından biri olan NSGA-III’te, NSGA-II’deki kalabalık uzaklık operatörü bir kümeleme operatörü ile değiştirilip çeşitliliği garanti etmek için iyi dağılmış referans noktaları kümesi kullanılır.

Genellikle, dört veya daha fazla amaç fonksiyonu içeren ve bazen 10-15 amaç fonksiyonuna sahip olmayı gerektiren birçok optimizasyon probleminde çok sayıda amacı ele almada optimizasyon algoritmaları için bir dizi zorluk meydana gelmektedir. Rastgele seçilen amaç fonksiyonları vektöründeki baskın çözümlerin oranı, amaç fonksiyonu sayısı arttıkça katlanarak artar. Baskın çözümler popülasyon alanlarının çoğunu kapladığından seçkinleri koruyan herhangi bir evrimsel MOO algoritması, popülasyonda yeterli sayıda yeni çözüme uyum sağlamakta güçlüklerle karşılaşır. Bu sorun, arama sürecini önemli derecede yavaşlatır. Çeşitlilik koruma operatörünün uygulanması hesaplama bakımından maliyetlidir. Ayrıca büyük boyutlu bir yüzeyin görselleştirilmesi de zordur. Bu zorluklar nedeniyle çok amaçlı optimizasyon güçleşir.

EA tabanlı MOO algoritmalarında karşılaşılan bazı zorluklar aşağıdaki gibi verilebilir. *Popülasyonun büyük bir kısmı baskındır.* Amaç fonksiyonu sayısındaki artışla, rastgele oluşturulmuş bir popülasyonun giderek daha büyük bir kısmı baskın hale gelir. Çoğu evrimsel çok amaçlı optimizasyon algoritması, bir popülasyondaki baskın çözümleri vurguladığından, çok amaçlı optimizasyon problemlerinin ele alınmasında, bir nesilde

yeni çözümler oluşturmak için fazla yer yoktur. Bu durum, arama sürecini yavaşlatır ve evrimsel çok amaçlı optimizasyon algoritması yetersiz hale gelir.

Çeşitlilik ölçüsünün değerlendirilmesi, hesaplama bakımından maliyetlidir. Bir popülasyondaki çözümlerin kalabalıklaşma derecesini belirlemek için komşuların tanımlanması, büyük boyutlu bir alanda hesaplama açısından zordur. Hesaplamaları daha hızlı hale getirmeye çalışılırken çeşitlilik tahminindeki herhangi bir ödünleşme kabul edilemez bir çözüm dağılımına neden olabilir.

Çaprazlama işlemi verimsiz olabilir. MOO problemlerinde, büyük boyutlu bir uzayda az sayıda çözüm elde edilmesi isteniyorsa çözümlerin birbirinden oldukça uzak olması olasıdır. Böyle bir popülasyonda, çaprazlama işleminin etkisi sorgulanabilir. Birbirinden uzak iki ebeveyn çözümün ebeveynlerden de uzak yavru çözümler üretmesi muhtemeldir. Bu nedenle, çok amaçlı problemleri verimli bir şekilde ele almak için özel çaprazlama operatörleri (çiftleşme kısıtlaması) gerekli olabilir.

Ödünleşme yüzeyinin temsili zordur. Daha yüksek boyutlu bir ödünleşme yüzeyini temsil etmek için üstel olarak daha fazla nokta gereklidir. Bu nedenle, Pareto yüzeyi temsil etmek için büyük bir popülasyon boyutu gerekir. Bu da karar vericinin tercih edilen çözümü anlamasını ve uzlaşık bir karar vermesini zorlaştırmaktadır.

Performans ölçütlerini hesaplamak maliyetlidir. Bir algoritmanın diğerine karşı performansını belirlemek için daha yüksek boyutlu nokta kümelerinin birbiriyle karşılaştırılması gerektiğinden, daha büyük bir hesaplama çabası gerekir.

Görselleştirme zordur. Daha yüksek boyutlu bir ödünleşme yüzeyinin görselleştirilmesi çok amaçlı problemler için zor olabilir.

NSGA-III yöntemi, temelde seçim operatöründeki önemli değişiklikler dışında NSGA-II yöntemine benzerdir. NSGA-II'den farklı olarak NSGA-III'te, popülasyon üyeleri arasındaki çeşitliliğin korunmasına, bir dizi iyi yayılmış referans noktasının sağlanması

ve uyarlanabilir şekilde güncellenmesi yardımcı olur. NSGA-II'deki kalabalık uzaklık operatörü NSGA-III'te, *popülasyonun baskınlık düzeylerine göre sınıflandırılması*, *hiperdüzlemde referans noktalarının belirlenmesi*, *popülasyon üyelerinin uyarlanabilir normalizasyonu*, *ilişkilendirme işlemleri* ve *niş koruması işlemi* ile değiştirilir. Bu işlemlere aşağıda detaylı biçimde yer verilmiştir.

(i) Popülasyonun Baskınlık Düzeylerine Göre Sınıflandırılması: Baskınlık ilkesini kullanarak baskın yüzeyleri belirlemeye yönelik şekil 4.7'deki işleyiş biçimi NSGA-III'te de kullanılır. Baskınlık seviyesi 1'den l 'ye kadar olan tüm popülasyon üyeleri ilk olarak S_t 'ye dahil edilir. Eğer $|S_t| = N$ ise bir sonraki nesil $P_{t+1} = S_t$ ile başlatılır. $|S_t| > N$ için, 1 ile $(l-1)$ yüzeylerinden $P_{t+1} = \bigcup_{i=1}^{l-1} F_i$ popülasyon üyeleri seçilir ve kalan $K = N - |P_{t+1}|$ popülasyon üyeleri son yüzey F_l 'den seçilir.

(ii) Hiperdüzlem Üzerinde Referans Noktalarının Belirlenmesi: NSGA-III'te referans noktaları önceden tanımlanabilir ya da kullanıcı tarafından tercihen belirlenebilir. Herhangi bir tercih bilgisi yoksa, önceden tanımlanmış referans noktaları kullanılabilir. Deb ve Jain (2014) çalışmalarında Das ve Dennis'in (1998) tüm amaç eksenlerine eşit eğimli ve her ekseninde bir kesişme noktası olan normalleştirilmiş bir hiperdüzlem üzerine noktaları yerleştiren sistematik yaklaşımını kullanmışlardır. Her bir amaç boyunca p bölüm dikkate alınırsa, M amaçlı bir problemdeki toplam referans noktası sayısı H

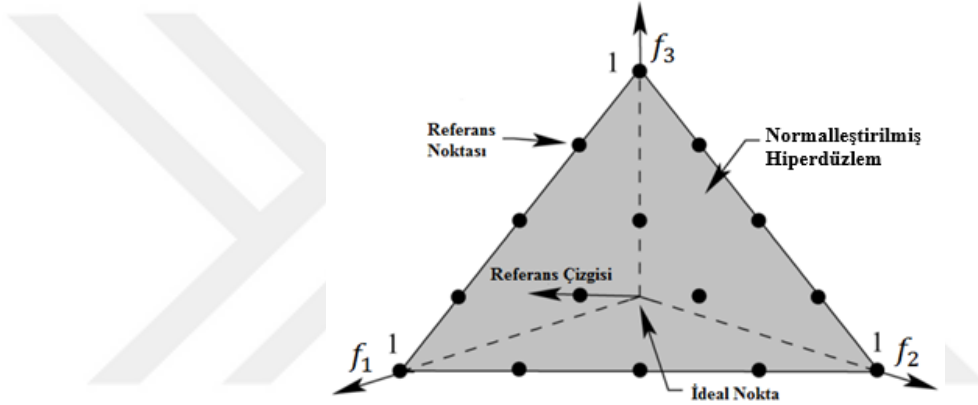
$$H = \binom{M + p - 1}{p} \quad (4.19)$$

ile hesaplanır. Burada H referans noktalarının sayısı, amaç sayısına ve referans noktalarının aynı yüzey üzerinde birbirlerine olan uzaklığını belirleyen p bölünmesine bağlıdır. Elde edilen çözümlerde çeşitliliğin sağlanması amacıyla bir referans çizgisinde kaç referans noktasının kullanılacağını belirlemek için bölüm sayısından yararlanılır.

Örneğin, üç amaçlı ($M = 3$) bir problemde, referans noktaları tepe noktası $(1,0,0)$, $(0,1,0)$ ve $(0,0,1)$ olan bir üçgen üzerinde oluşturulsun. Her bir amaç eksenini için dört bölüm

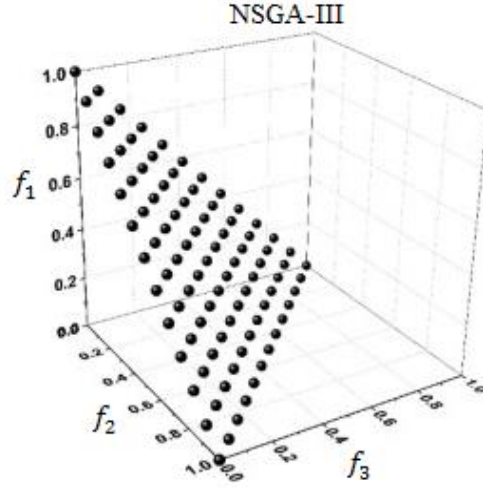
($p = 4$) seçilirse, oluşturulacak olan referans noktası sayısı $H = \binom{3+4-1}{4} = 15$ olarak

bulunur. Bu referans noktaları Şekil 4.9'da gösterilmiştir. Şekil 4.9'da görüldüğü gibi referans noktaları, normalleştirilmiş hiperdüzlemin tamamında yaygın ve düzgün bir şekilde dağılmıştır. Elde edilen çözümlerin de Pareto yüzey üzerinde veya yakınında yaygın olarak dağılması olasıdır. Dolayısıyla bu süreç, yeni popülasyonun çeşitliliğini ve dağılımını korumak amacıyla çeşitli ve iyi dağıtılmış baskın çözümlerin seçimini teşvik eder (Yannibelli vd. 2020). Böylece, bu yöntemin karar verme ve MOO'nun uygulaması bakımından daha fazla kullanımı sağlanır.



Şekil 4.9 $M = 3$ ve $p = 4$ iken bir problem için normalleştirilmiş bir referans düzleminde gösterilen 15 yapılandırılmış referans noktası (Deb ve Jain 2014)

(iii) Popülasyon Üyelerinin Uyarlanabilir Normalizasyonu: NSGA-III, belirli bir yüzeydeki çözümlerinin çeşitliliğini korumak için referans noktalarını kullanır. NSGA-III'te, farklı büyüklük aralığındaki amaçları ele almak için başarı ölçeklendirme fonksiyonu (Achievement Scalarization Function-ASF) tabanlı bir normalizasyon benimsenmiştir. Yüzey çözümleri normalleştirilmiş bir düzlem üzerine yansıtılır. Üç amaçlı bir optimizasyon problemi için normalleştirilmiş bir düzlem, Şekil 4.10'daki gibi eşit aralıklı referans noktalarına bölünmüştür.

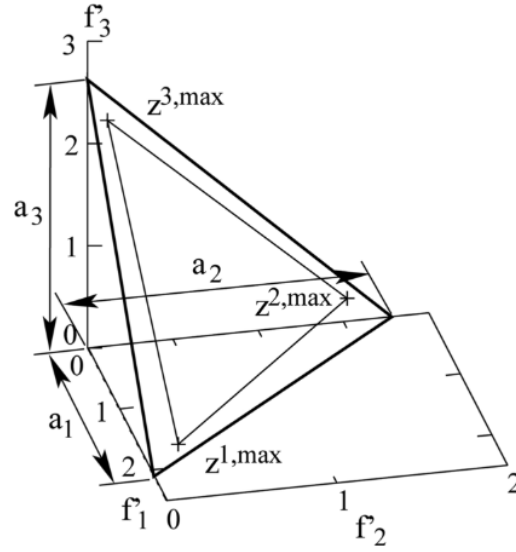


Şekil 4.10 Referans düzlemi (Chaudhari vd. 2022)

İlk olarak S_t popülasyonunun ideal noktası için $\bigcup_{\tau=0}^t S_\tau$ 'deki her bir amaç fonksiyonunun $i = 1, 2, \dots, M$ minimum değeri (z_i^{min}) belirlenir ve $\bar{z} = (z_1^{min}, z_2^{min}, \dots, z_M^{min})$ oluşturulur. S_t 'nin her bir amaç değeri, amaç fonksiyon değeri f_i 'den z_i^{min} çıkarılarak elde edilir. Elde edilen S_t 'nin ideal noktası bir sıfır vektörü olur.

$$f'_i(x) = f_i(x) - z_i^{min}, \quad i = 1, 2, \dots, M \quad (4.20)$$

Her bir i . amaç eksenindeki uç nokta ($z^{i,max}$), karşılık gelen başarı ölçeklendirme fonksiyonunu minimum yapan çözüm ($x \in S_t$) bulunarak belirlenir. Bu M uç vektörleri daha sonra M -boyutlu bir hiperdüzlem oluşturmak için kullanılır. Böylece, i . amaç ekseninin α_i kesişimi ve doğrusal hiperdüzlemi hesaplanabilir. Bozulmuş (dejenere) durumları ve negatif olmayan kesişimleri ele alırken dikkatli olunmalıdır. Şekil 4.11'de üç amaçlı bir problem için kesişim noktalarının hesaplanması ve ardından uç noktalardan hiperdüzlem oluşturma işlemi yer almaktadır.



Şekil 4.11 Üç amaçlı problem için kesişim noktalarının hesaplanması ve ardından uç noktalardan hiperdüzlem oluşturma işlemi (Deb ve Jain 2014)

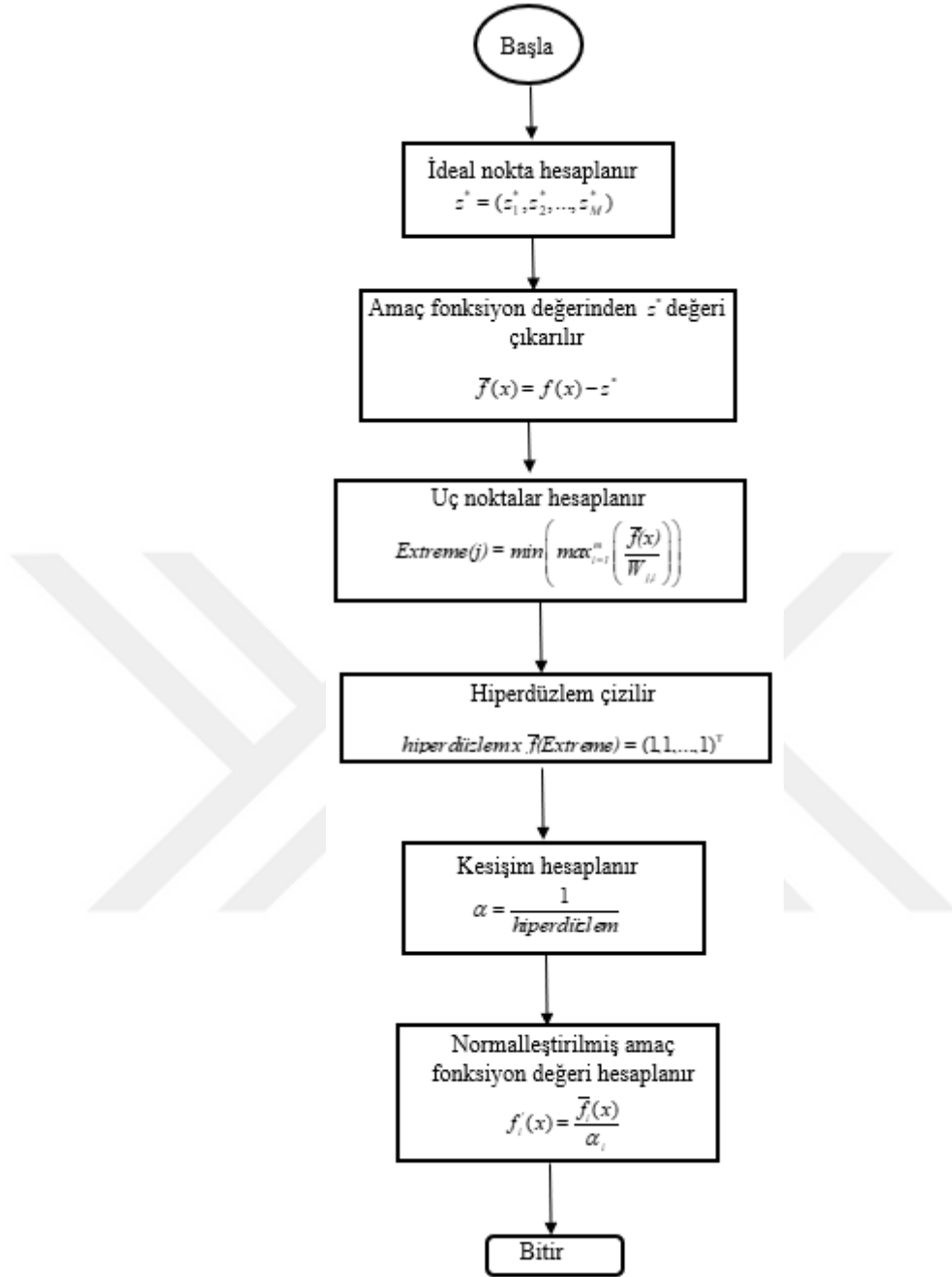
Amaç fonksiyonları

$$f_i^{nor}(x) = \frac{f_i'(x)}{\alpha_i}, \quad i = 1, 2, \dots, M \quad (4.21)$$

eşitliği ile normalleştirilir. Her bir normalleştirilmiş amaç eksenindeki kesişme noktası $f_i^{nor}(x) = 1$ noktasındadır. Bu kesişme noktalarıyla oluşturulan hiperdüzlem

$$\sum_{i=1}^M f_i^{nor}(x) = 1 \text{ olur. Das ve Dennis'in (1998) sistematik yaklaşımı kullanılarak}$$

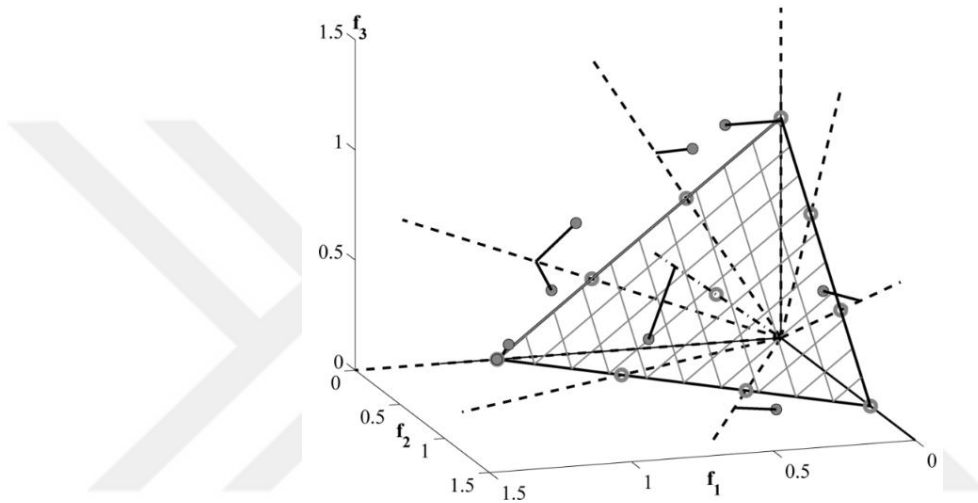
hesaplanan referans noktaları normalleştirilmiş hiperdüzlem üzerinde yer alır. Referans noktalarının kullanıcı tarafından tercih edilmesi durumunda, referans noktaları eşitlik (4.21) ile şekil 4.12'de oluşturulan normalleştirilmiş hiperdüzlem üzerinde eşleştirilir. Normalleştirme prosedürü ve hiperdüzlemin oluşturulması her nesilde simülasyonun başlangıcından itibaren bulunan uç noktalar kullanılarak yapıldığından, NSGA-III prosedürü her nesilde S_t üyeleri tarafından yayılan uzayda uyarlamalı olarak çeşitliliği korur. Bu, NSGA-III'ün amaç değerleri farklı biçimde ölçeklendirilebilen Pareto yüzey problemlerini çözmeyi sağlar. NSGA-III'ün normalleştirme işlemleri şekil 4.12'de verilmiştir.



Şekil 4.12 NSGA-III'ün normalleştirme işlemleri (Gu vd. 2021)

(iv) **İlişkilendirme İşlemi:** Her bir amaç fonksiyonu, amaç uzayında S_t , $t = 1, 2, \dots, n_{\text{gen}}$, üyelerinin boyutuna göre uyarlamalı olarak normalleştirildikten sonra her bir popülasyon üyesi bir referans noktasıyla ilişkilendirilir. Referans noktası orijinle birleştirilerek hiperdüzlemdeki her bir referans noktasına karşılık gelen bir referans çizgisi tanımlanır. S_t 'nin her bir popülasyon üyesinin tüm referans çizgilerine olan dik uzaklığı hesaplanır.

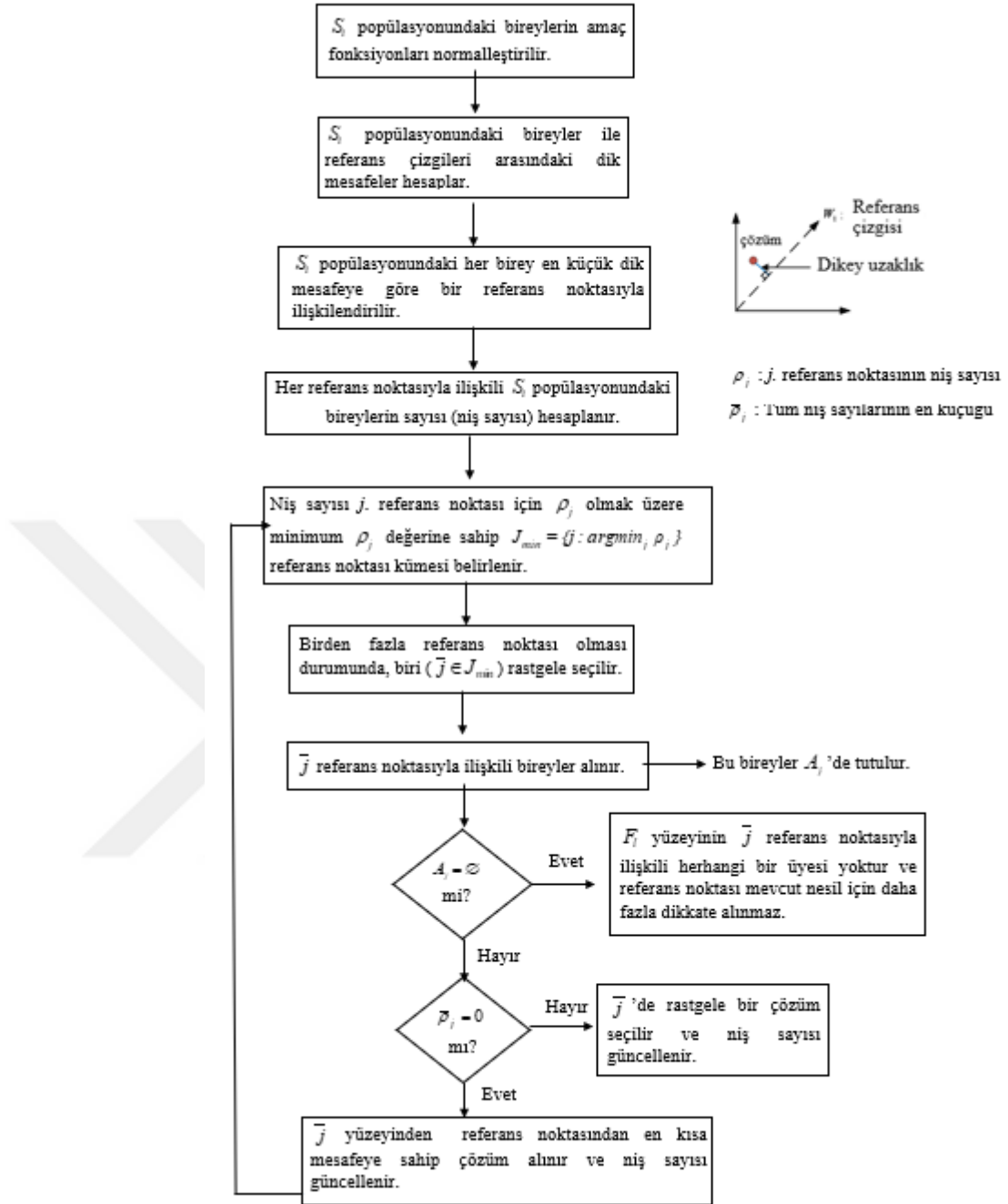
Normalleştirilmiş amaç uzayında, referans çizgisi bir popülasyon üyesine en yakın olan referans noktasının popülasyon üyesiyle ilişkili olarak kabul edilir. Şekil 4.13'te popülasyon üyelerinin referans noktaları ile ilişkisi görülmektedir. Normalleştirilmiş düzlem üzerine yerleştirilen her çözüm en yakın referans noktası ile ilişkilendirildikten sonra, seçilen çözüm kümesinde maksimum sayıda referans noktası temsil edilecek şekilde bir seçim prosedürü izlenir (Chaudhari vd. 2022). NSGA-III'te, baskın çözümleri vurgulamanın yanı sıra referans noktalarının her biriyle ilişkili olan popülasyon üyeleri de vurgulanır.



Şekil 4.13 Popülasyon üyelerinin referans noktalarıyla ilişkisi (Deb ve Jain 2014)

(v) Niş Koruması İşlemi: Bir referans noktası kendisiyle ilişkili bir veya daha fazla popülasyon üyesine sahip olabileceği gibi kendisiyle ilişkili herhangi bir popülasyon üyesine sahip olmayabilir. $P_{t+1} = S_t / F_l$ 'den her bir referans noktasıyla ilişkili popülasyon üyeleri sayısı hesaplanır. F_l yüzeyinden P_{t+1} 'de en az temsil edilen referans noktalarıyla ilişkili popülasyon üyelerini seçmek için niş koruması uygulanır. Niş koruması mümkün olduğunca çok sayıda sağlanan referans noktası için bir popülasyon üyesi seçmeye vurgu yapar. Seçim işleminde önce niş sayısı az olan referans noktalarına atanmış çözümler tercih edilir. Bunun nedeni, niş sayısı az olan referans noktalarına atanmış çözümlerin seçimi ile daha iyi bir çeşitliliğe sahip olunabileceği düşüncesidir. Az sayıda çözümle temsil edilen referans noktalarının Pareto yüzeyinin daha az keşfedilmiş veya daha az yoğun bölgelerini temsil ettiği ve bu nedenle de daha fazla dikkat edilmesi gerektiği düşünülür.

Niş sayısı, j . referans noktası için ρ_j olarak gösterilmiştir, $j = 1, 2, \dots, H$. Deb ve Jain'in (2014) tasarladığı yeni niş koruma işleminde ilk olarak minimum ρ_j değerine sahip $J_{\min} = \{j : \operatorname{argmin} \rho_j\}$ referans noktası kümesi belirlenir. Minimum ρ_j değerine sahip birden fazla referans noktası olması durumunda, referans noktalarından biri rastgele seçilir, $\bar{j} \in J_{\min}$. Eğer, $\rho_j = 0$ ise, F_l kümesinde $\bar{j}$ ile ilgili iki senaryo olabilir. Birincisi, F_l yüzeyinde $\bar{j}$ referans noktasıyla ilişkili bir veya daha fazla üye vardır. Bazı durumlarda bir referans noktasına atanmış bir veya daha fazla çözüm olduğunda ($\rho_j \neq 0$), bu çözümlerin sayısı yeterli olmayabilir. Pareto yüzeyin belirli bir bölgesinde çözümlerin yeterince yoğun veya çeşitli olmaması durumunda niş sayısı hala sıfır olabilir. Bu durumda, referans çizgisinden en kısa dik mesafeye sahip olan üye P_{t+1} 'e eklenir ve $\bar{j}$ referans noktası için ρ_j sayısı bir artırılır. İkincisi, F_l yüzeyinin $\bar{j}$ referans noktasıyla ilişkili herhangi bir üyesi yoktur. Bu durumda, niş sayısı sıfır olur ve referans noktası mevcut nesil için daha fazla dikkate alınmaz. $\rho_{\bar{j}} \geq 1$ olması durumunda (S_t/F_l 'de referans noktasıyla ilişkili bir üye mevcut ise), F_l yüzeyinden $\bar{j}$ referans noktasıyla ilişkili rastgele seçilen bir üye P_{t+1} 'e eklenir, $\rho_{\bar{j}}$ sayısı bir artırılır. Niş sayıları güncellendikten sonra, P_{t+1} 'in tüm boş popülasyon yuvalarını doldurmak için prosedür toplam K kez tekrarlanır. Bu prosedürle, çözümlerin çeşitliliği korunarak ve referans noktaları arasında daha dengeli bir dağılım sağlanarak Pareto yüzeyinde daha iyi bir yayılım elde edilmesi amaçlanır. Şekil 4.14'te F_l yüzeyinden $K = N - |P_{t+1}|$ bireyin seçimi yer almaktadır.



Şekil 4.14 F_i yüzeyinden $K = N - |P_{t+1}|$ bireyin seçimi (Gu vd.2021)

(vi) **Yavru Popülasyonu Oluşturmak için Genetik İşlemler:** P_{t+1} oluşturulduktan sonra, genetik operatörler uygulanarak yavru popülasyon Q_{t+1} oluşturulur. NSGA-III'te elitist çözüm seçimi gerçekleştirilir. Her referans noktasının referans çizgisine en yakın çözümlerinin vurgulanmasıyla çözümler arasındaki çeşitlilik sağlanarak popülasyon

korunmaya çalışılır. Deb ve Jain (2014) referans noktası sayısını popülasyon büyüklüğüne yaklaşık olarak eşit alarak her referans noktasına karşılık gelen Pareto optimal yüzeye yakın bir popülasyon üyesi bulmayı amaçlamışlardır. Bhesdadiya vd. (2016) ise popülasyon büyüklüğünün referans noktası sayısının dört katı büyüklüğünde seçilebileceğini belirtmiştir.

NSGA-III'te her bir referans noktası için yalnızca bir popülasyon üyesinin bulunması beklendiğinden genetik operatörlerden seçim operatörü kullanılmaz. Çünkü seçim operatörünün kullanımı, farklı referans noktaları arasında rekabetin kurulmasına izin verir (Seada ve Deb 2015). Q_{t+1} yavru popülasyonu, P_{t+1} 'den rastgele ebeveynler seçerek çaprazlama ve mutasyon operatörleri uygulanarak oluşturulur (Deb ve Jain 2014). Ebeveyn çözümlerine daha yakın yavru çözümler oluşturmak için, SBX operatöründe nispeten daha büyük bir dağılım indeksi değeri kullanılmalıdır.

(vii) NSGA-III'ün parametresizlik özelliği: NSGA-II'de olduğu gibi, NSGA-III algoritması da popülasyon büyüklüğü, sonlandırma parametresi, çaprazlama ve mutasyon olasılıkları ve bunlarla ilişkili parametreler gibi olağan GA parametreleri dışında herhangi bir yeni parametre ayarlaması gerektirmez. Referans noktalarının sayısı, doğrudan istenen ödünleşim noktası sayısı ile ilgili olduğundan algoritmik bir parametre değildir. Popülasyon büyüklüğü, $N \approx H$ olarak H 'ye bağlıdır. Referans noktalarının konumu da kullanıcının elde edilen çözümlerde ulaşmak istediği tercih bilgisine bağlıdır.

Akış şeması şekil 4.15'te verilen NSGA-III yönteminin algoritmik adımları aşağıda tanımlanmıştır.

NSGA-III Algoritması:

Adım 0: t , yineleme sayısı olmak üzere, NSGA-III'ün ayarlanabilir parametreleri, popülasyon büyüklüğü (N), girdi değişken (karar değişkeni) sayısı (p), çaprazlama olasılığı (P_c), çaprazlama indeksi (η_c), mutasyon olasılığı (P_m), mutasyon indeksi (η_m), referans bölüm sayısı (n_{div}) ve maksimum yineleme sayısı (n_{gen}), uzman görüşüne göre

belirlenir, $t = 1, 2, \dots, n_{gen}$. Karar değişkenleri için belirlenen tanım aralıklarından karar değişkenleri değerleri tanımlanır.

Adım 1: Başlangıç popülasyonu ve düzgün dağılım ile referans noktaları (H) oluşturulur.

Adım 2: Bireyler uygunluk fonksiyon değerlerine göre sıralanır.

Adım 3: Genetik operatör uygulamaları çaprazlama ve mutasyon aşamalarından oluşur. *Adım 0*'da belirtilen ayarlanabilir parametre değerleri kullanılarak genetik operatör işlemleri uygulanır.

- Baskınlık kriterine göre sıralanan bireylerden rastgele ebeveynler seçilir.
- Seçilen ebeveynlere SBX operatörü uygulanarak yavrular üretilir.
- SBX operatörü ile üretilen bazı yavrulara Polinomsal mutasyon operatörü uygulanır.

Adım 4: $t = 1, 2, \dots, n_{gen}$ olmak üzere, N ve $2N$ bireyli yeni popülasyon oluşturulur.

- $Q_t(N)$ yavru popülasyonu oluşturulur.
- Mevcut popülasyon ile oluşturulan yavru popülasyon birleştirilir, $P_t(N) \cup Q_t(N) = R_t(2N)$. Oluşturulan $R_t(2N)$ popülasyonundaki tüm çözümlerin birlikte değerlendirilmesiyle seçkinlik işlemi gerçekleştirilir.

Adım 5: Baskın sıralama yöntemi kullanılarak birleşik popülasyon $R_t(2N)$ farklı baskınlık düzeylerine ($F_1, F_2, \dots, F_l$) ayrılır, $t = 1, 2, \dots, n_{gen}$.

Adım 6: $F_1, F_2, \dots, F_l$ yüzeylerinden bireyler seçilir. Yüzeylerden seçilen bireylerle S_t yeni popülasyonu oluşturulur. S_t popülasyonunda birey sayısı ilk kez N 'ye eşit veya

N 'den büyük olana kadar F yüzeylerinden birey seçimi devam eder, $S_t = \bigcup_{i=1}^l F_i$,

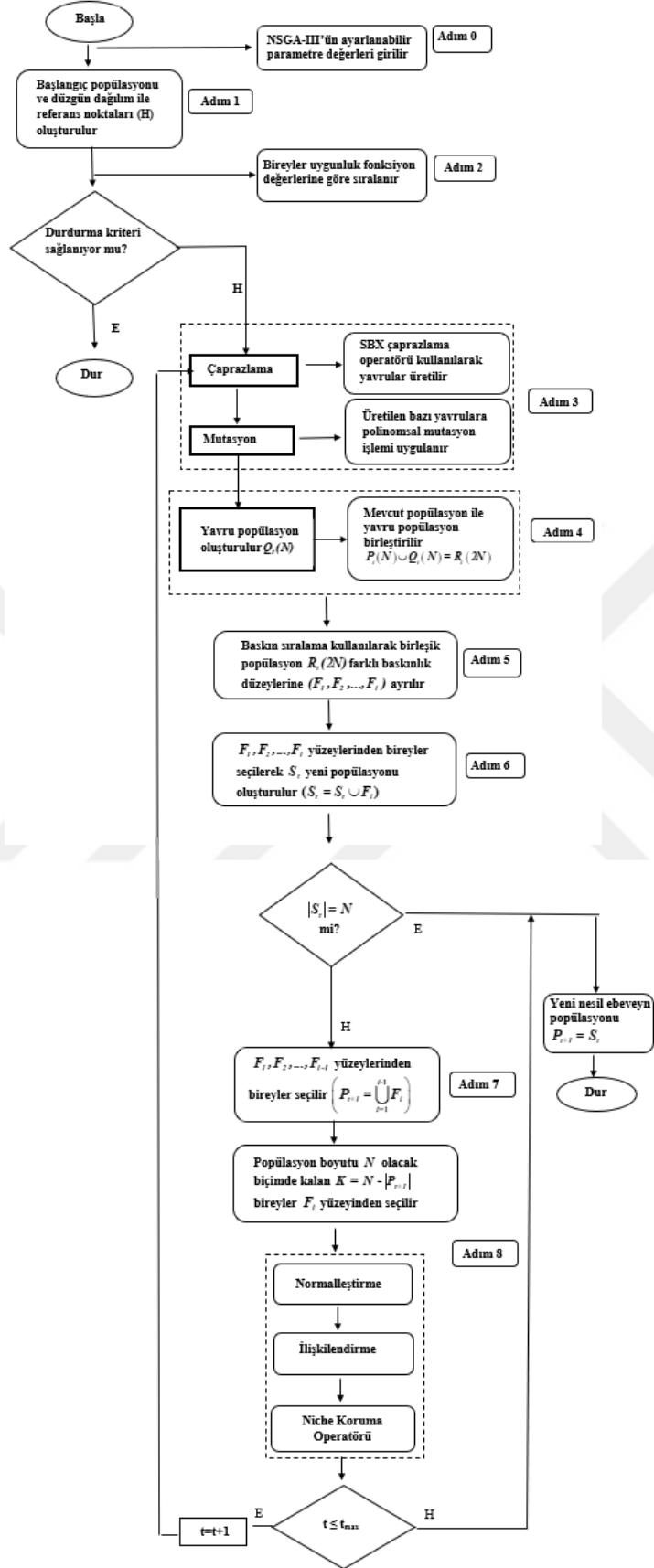
$t = 1, 2, \dots, n_{gen}$.

Adım 7: S_t popülasyonunun boyutunun N 'ye eşit olup olmadığı kontrol edilir. Popülasyonun birey sayısı N 'ye ulaşmışsa, $S_t(N)$ popülasyonu bir sonraki yineleme için $P_{t+1}(N)$ ebeveyn popülasyonu olarak ele alınır, $P_{t+1}(N) = S_t(N)$ ve algoritma durur. Aksi halde, algoritma devam eder, $F_1, F_2, \dots, F_{l-1}$ yüzeylerinden bireyler seçilir,

$$P_{t+1} = \bigcup_{i=1}^{l-1} F_i, t = 1, 2, \dots, n_{gen}.$$

Adım 8: Popülasyon boyutu N olacak biçimde kalan $K = N - |P_{t+1}|$ birey F_l yüzeyinden seçilir. $t = 1, 2, \dots, n_{gen}$ olmak üzere, K bireyin seçimi için aşağıdaki işlemler uygulanır.

- İlk olarak S_t popülasyonundaki bireylerin amaç fonksiyon değerleri normalleştirilir.
- Hiperdüzlemdeki referans çizgileri tanımlanır. S_t popülasyonundaki bireyler ile referans çizgileri arasındaki dik mesafeler hesaplar. S_t popülasyonundaki her birey en küçük dik mesafeye göre bir referans noktasıyla ilişkilendirilir.
- Her referans noktası için, referans noktasıyla ilişkili S_t popülasyonundaki bireylerin sayısı olarak tanımlanan *niş sayısı* hesaplanır. Hesaplanan niş sayısına göre K birey seçilir. Başlangıçta belirlenen yineleme sayısına ulaşılmış ise algoritma sonlandırılır. Aksi halde, *Adım 3*'e gidilir. Şekil 4.15'te NSGA-III'ün akış şeması yer almaktadır.



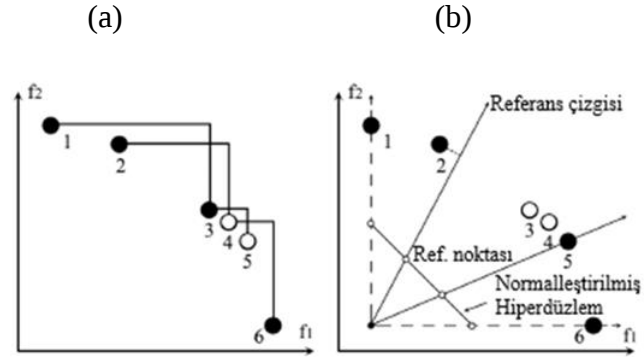
Şekil 4.15 NSGA-III'ün akış şeması (Merah ve Adjabi 2018)

4.4.3 NSGA-II ile NSGA-III yöntemlerinin karşılaştırılması

NSGA-II ve NSGA-III yönteminin Pareto çözüm kümesini elde etme arayışı benzerdir. Her iki yöntem de rastgele bir başlangıç popülasyonu ile aramaya başlar. Her nesilde bu yöntemler, mevcut popülasyondan yeni bir popülasyon oluşturmak için aynı çaprazlama ve mutasyon operatörlerini kullanır. Fakat, NSGA-III algoritması, Q_t yavru popülasyonu oluşturma sürecinde P_t popülasyonu üzerinde herhangi bir doğrudan seçim operatörü kullanmaz.

Bu algoritmalar, birleştirilmiş mevcut ve yavru popülasyondan hangi çözümlerin gelecek nesil için yeni popülasyonu oluşturacağını belirlemede kullanılan seçim sürecine göre farklılık gösterir. NSGA-II’de seçim süreci, öncelikle birleştirilmiş popülasyondaki çözümlerin baskınlık düzeyini ve ardından bu çözümlerin kalabalık uzaklığını dikkate alır. Kalabalık uzaklık, bir çözümün komşu çözümlere olan mesafesini temsil eder ve daha büyük kalabalık uzaklığa sahip baskın çözümlerin seçimini vurgular. Ancak bu süreç, iyi dağılmış, baskın çözümlerin seçimini garanti etmez (Yannibelli vd. 2020). Çok amaçlı problemler için kalabalık uzaklık ölçütü iyi performans göstermemektedir. Bu nedenle, NSGA-III’teki seçim mekanizması, verilen referans noktalarına göre S_t ’deki üyelerin daha sistematik bir analizini yapacak biçimde değiştirilmiştir (Yuan vd. 2014). NSGA-III algoritması, iyi yayılmış referans noktaları ve bu noktaların uyarlamalı olarak güncellenmesiyle, NSGA-II’nin popülasyon üyeleri arasındaki çeşitliliği sürdürme yetersizliğini ortadan kaldırır (Bheshadiya vd. 2016). İyi yayılmış referans noktalarına dayanan bu seçim sürecini kullanarak NSGA-III yöntemi, baskın çözümlerin hem çeşitliliği hem de dağılımı bakımından daha iyi Pareto kümesine ulaşma olasılığına sahiptir. Şekil 4.15’te NSGA-II ve NSGA-III yöntemlerinin çözümler arasındaki çeşitliliği korumak için kullandığı operatörler yer almaktadır. NSGA-III, çözümler arasındaki çeşitliliği korumak için bir dizi referans yönü kullanırken, NSGA-II aynı amaç için kalabalık mesafe operatörünü kullanır. Şekil 4.16’da gösterildiği gibi iki algoritmanın yeni nesil için mevcut baskın altı noktadan iki noktayı seçmesi gerektiği varsayılırsa, NSGA-II, şekil 4.16 (a)’daki siyah noktaları beyaz noktalara tercih eder. Çünkü bu noktalar arasındaki kalabalık uzaklıklar daha büyüktür. Öte yandan, NSGA-III, sağlanan

dört referans yönünü en iyi temsil eden dört noktayı, şekil 4.16 (b)'de gösterildiği gibi tercih eder.



Şekil 4.16 (a) NSGA-II ve (b) NSGA-III yöntemlerinin çözümler arasındaki çeşitliliği korumak için kullandığı operatörler (Seada ve Deb 2015)

5. KARMA OPTİMİZASYON PROBLEMLERİ İÇİN UYARLANMIŞ GENETİK ALGORİTMA

Literatürde mevcut olan optimizasyon yöntemlerinin çoğunda, karar değişkenlerinin sürekli değerli olarak ele alındığı görülmektedir. Gerçek dünya problemlerinin matematiksel modelindeki karar değişkenleri kesikli değerli olduğunda optimizasyon yöntemlerinin uyarlanması gerekir. Tam sayılı ve rasyonel değerler aldığı göz önünde bulundurulmuş kesikli değişken değerleri için yuvarlama yöntemleri kullanılarak tüm değişkenler sürekli olarak ele alınıp optimizasyon problemi çözülebilir. Fakat, bu durumda optimizasyon uygun olmayan çözümlerle sonuçlanabilir. Bunu engellemek için karar değişkenlerinin kesikli olma durumu, problemin çözümünde bilgi kaybına neden olmayacak şekilde dikkate alınmalıdır. Bu bölümde karma verili problemlerin optimizasyonu için GA yönteminin işleyiş adımlarının uyarlanmasına yer verilecektir.

5.1 Uyarlanmış Genetik Algoritma

Karma veri içeren problemlerin klasik optimizasyon yöntemleriyle çözümü, düşük verimlilik, sınırlı güvenilirlik ve yerel bir optimumda sıkışıp kalma gibi dezavantajlara sahiptir. Bu tür problemler için GA, uygun bir esnek hesaplama yöntemidir.

Bu tez çalışmasında, değer kodlama kullanılarak NSGA-II ve NSGA-III yöntemlerine değişken gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin uygulanması aşamalarında uyarlamalar yapılarak karma verili MOO problemleri için çözüm yöntemleri önerilmiştir. Değer kodlamayla karma verili MOO problemlerinin optimizasyonu için önerilen yöntemler, MDNSGA-II ve MDNSGA-III olarak adlandırılmıştır.

5.1.1 Uyarlanmış karar değişkenlerinin gösterimi

Karar değişkenlerinin sürekli ve kesikli değer alması durumunda GA'da bilinen değişken gösterimi uygulanamaz. Bu durumda karar değişkenlerinin ikili kodlama (Wu ve Chow 1994, Lin ve Hajela 1992) veya değer kodlamasıyla gösterimi gerekir. Bu tez

çalışmasında hesaplama süresini kısaltması ve karma verili problemler için kolay uygulanabilir olması nedeniyle değer kodlama kullanılmıştır. Aşağıda, tezin özgünlüğünü oluşturan MDNSGA-II ve MDNSGA-III yöntemlerinin elde edilmesi için uygulanan uyarlamalar detaylı biçimde açıklanmıştır.

Değer kodlamasında sürekli ve tam sayılı değişkenlerin kendi değeriyle çalışılırken, karar değişkenlerinin rasyonel reel değer alması durumunda değişken gösteriminde uyarlama yapılması gerekir. Karar değişkeninin tanım kümesindeki her bir rasyonel değer küçükten büyüğe sıralanır. Sıralanan değerler sıra numarası ile dizilere eşlenir. Böylece rasyonel değerli kesikli değişkenin her bir değeri bir sıra numarasıyla ifade edilmiş olur.

Örneğin, rasyonel değerli bir değişkenin tanım kümesi $S \subset \mathbb{Q}$, $S = \{s_1, s_2, \dots, s_q\}$ olmak üzere, rasyonel değerli kesikli değişkenin her bir değerinin I tam sayılar kümesinde pozitif bir tam sayılı diziye eşlenmesi

$$\begin{array}{ccccccc} S = \{s_1, s_2, s_3, \dots, s_q\} & & & & & & \\ \downarrow & \downarrow & \downarrow & \dots & \downarrow & & \\ I = \{1, 2, 3, \dots, q\} & & & & & & \end{array} \quad (5.1)$$

biçiminde gösterilir. (5.1) ifadesinde belirtildiği gibi, s_1 için 1, s_2 için 2, ..., s_q için q sıra numarası kullanılır. Böylece, rasyonel değerli kesikli değişkenin aldığı değerler yerine değerlerin sıra numaraları kullanılır.

5.1.2 Uyarlanmış başlangıç popülasyonunun oluşturulması

GA'da başlangıç popülasyonu genellikle rastgele sayı üretici kullanılarak oluşturulur. Rastgele popülasyon üretimi GA'da başlangıç popülasyonunun oluşturulmasında yalnızca bir kez kullanılır. Bu adımda, belirlenen popülasyon büyüklüğünde kromozomlar (çözüm vektörleri) oluşturulur. Başlangıç popülasyonu rastgele üretildikten sonra her bir kromozom, amaç fonksiyonu aracılığıyla değerlendirilir ve buna göre bir uygunluk değeri atanır. Karar değişkenlerinin sürekli ve kesikli (tam sayılı ve/veya

rasyonel) değerler almasına göre başlangıç popülasyonundaki çözüm vektörlerinin rastgele üretilmesi farklılık gösterir.

Bir x karar değişkeni sürekli değerliyse ($x \in \mathbb{R}$), değer kodlaması kullanılarak $[lb, ub]$ aralığında eşitlik (4.1)'deki gibi rastgele çözüm üretilir. Karar değişkeninin kesikli değer alması tam sayı ve rasyonel sayı olmak üzere iki durumda mümkündür. x karar değişkeni tam sayı değerliyse ($x \in \mathbb{Z}$) değer kodlaması kullanılarak $[lb, ub]$ aralığında eşitlik (4.1)'e yuvarlama operatörünün dahil edilmesiyle

$$x = lb + \text{round}(a(ub - lb)) \quad (5.2)$$

biçiminde tam sayı değeri elde edilir. x kesikli karar değişkeni rasyonel değerli ise ($x \in \mathbb{Q}$) değişkenin aldığı değerlerin küçükten büyüğe sıralaması yapılarak sıra indisleri dikkate alınır.

Kesikli değişkenlerin aldığı rasyonel değerlerin sıra numaraları için rastgele sayı üretilir. r sıra numarası ve tam sayı eşlemesi yapılan kesikli değişkenin tanım aralığı $[1, q]$ olmak üzere kesikli değişken için üretilecek rastgele sayı

$$x^{(r)} = \text{round}(1 + a(q - 1)) \quad (5.3)$$

biçiminde elde edilir. $S = \{s_1, s_2, \dots, s_q\}$ kesikli değişken değerleri kümesi olmak üzere $\#(S) = q$ 'dur. $[1, q]$ aralığından q sayıda rastgele sıra numaraları üretilir ve üretilen sıra numaralarına göre, S kümesinde karşılık gelen değer seçilir. Oluşturulan yeni kesikli değer kümesiyle, kesikli değişkenler için başlangıç popülasyonu bileşen değerleri elde edilir.

5.1.3 Uyarlanmış genetik operatörlerin uygulanması

Karma optimizasyon problemlerinin çözümü için genetik operatörlerin uyarlanması işleyişinde popülasyondaki değişkenler tanım aralığında değerler alacağından seçim operatöründe herhangi bir uyarlama yapmaya gerek yoktur. MOO problemleri, NSGA-II

ve NSGA-III'ün çaprazlama ve mutasyon operatörlerinde yapılacak ayarlamalarla çözülebilir. Bu çalışmada, önerilen MDNSGA-II ve MDNSGA-III'te kullanılan SBX Çaprazlama ve Polinomsal Mutasyon operatörlerinde yapılan ayarlamalar aşağıda verilmiştir.

Sürekli değişkenler için SBX operatörüyle eşitlik (4.15)-(4.16) kullanılarak yavru çözümler oluşturulur. MDNSGA-II ve MDNSGA-III'te çaprazlama operatörü kullanılarak oluşturulan yavru çözümler, tam sayı ve rasyonel değerli kesikli değişkenlerin tanım kümesi dışında değerler aldığı anda, çaprazlama operatöründe yapılacak değişikliklerle değişkenin tanım kümesi içinde değer alması sağlanır. Tam sayılı değişkenler için yavru çözümler, çaprazlama operatörüne yuvarlama işlemi uygulanarak elde edilir. Yuvarlama

- En yakın tam sayıya yuvarlama
- %50 olasılıkla yukarı veya aşağı yuvarlama

yöntemlerinden biriyle yapılabilir (Rainer 2012). Tam sayılı değişkenin aldığı değerler ardışık olduğunda yavru çözümler

$$\begin{aligned} c_1 &= \text{round} \left(0.5[(y_1 + y_2) - \beta_q |y_2 - y_1|] \right) \\ c_2 &= \text{round} \left(0.5[(y_1 + y_2) + \beta_q |y_2 - y_1|] \right) \end{aligned} \quad (5.4)$$

biçiminde oluşturulur. Aksi halde, rasyonel değerli kesikli değişkenler için önerilen ayarlama yapılmalıdır. Ebeveyn çözümlere SBX operatörü uygulandığında, kesikli değişkenler için elde edilen yavru çözüm değerlerinin $S = \{s_1, s_2, \dots, s_q\}$ kümesindeki

elemanlardan biri olması istenir. Eğer, $c_1 \in [s_i, s_{i+1}]$, $i = 1, 2, \dots, q-1$ ise $s'_i = \frac{s_i + s_{i+1}}{2}$

hesaplanır. s'_i değeri eşitlik (4.16) ile elde edilen yavru çözüm değerleriyle karşılaştırılır.

Yapılan karşılaştırmada yavru çözüm değerinin s_i ya da s_{i+1} , $i = 1, 2, \dots, q-1$ değerlerinden hangisine daha yakın olduğu belirlenir. Yavru çözüm değerlerine yakın olan kesikli değişken değeri *kesikli yavru çözüm değeri* olarak alınır. Böylece, çaprazlama sonrasında elde edilen yavru çözüm

$$c_1 = \begin{cases} s_i & , \quad 0.5 \left[(y_1 + y_2) + \beta_q |y_2 - y_1| \right] \leq s'_i \\ s_{i+1} & , \quad d.y. \end{cases} \quad (5.5)$$

biçiminde elde edilir. c_2 yavru çözümü de eşitlik (5.5)'e benzer olarak hesaplanır.

SBX ile elde edilen yavru çözümlere mutasyon uygulanarak çözüm çeşitliliği sağlanır. Sürekli değişkenler için eşitlik (4.18) kullanılarak yeni çözüm elde edilir. MDNSGA-II ve MDNSGA-III'te Polinomsal mutasyon operatörü kullanılarak oluşturulan yeni çözüm, tam sayı ve rasyonel sayı değerli kesikli değişkenin tanım kümesi dışında değerler aldığı anda, mutasyon operatöründe yapılacak değişikliklerle değişkenin tanım kümesi içinde değer alması sağlanır. Tam sayılı değişkenler için yeni çözüm, çaprazlamada yapılan uyarlamalara benzer olarak mutasyon operatörüne uygulanarak elde edilir. Mevcut çözüm y 'den, yeni çözüm

$$c = \text{round} \left(y + \delta_q (ub - lb) \right) \quad (5.6)$$

biçiminde elde edilir. Kesikli değişkenler için $c_1 \in [s_i, s_{i+1}]$, $i = 1, 2, \dots, q-1$ ise

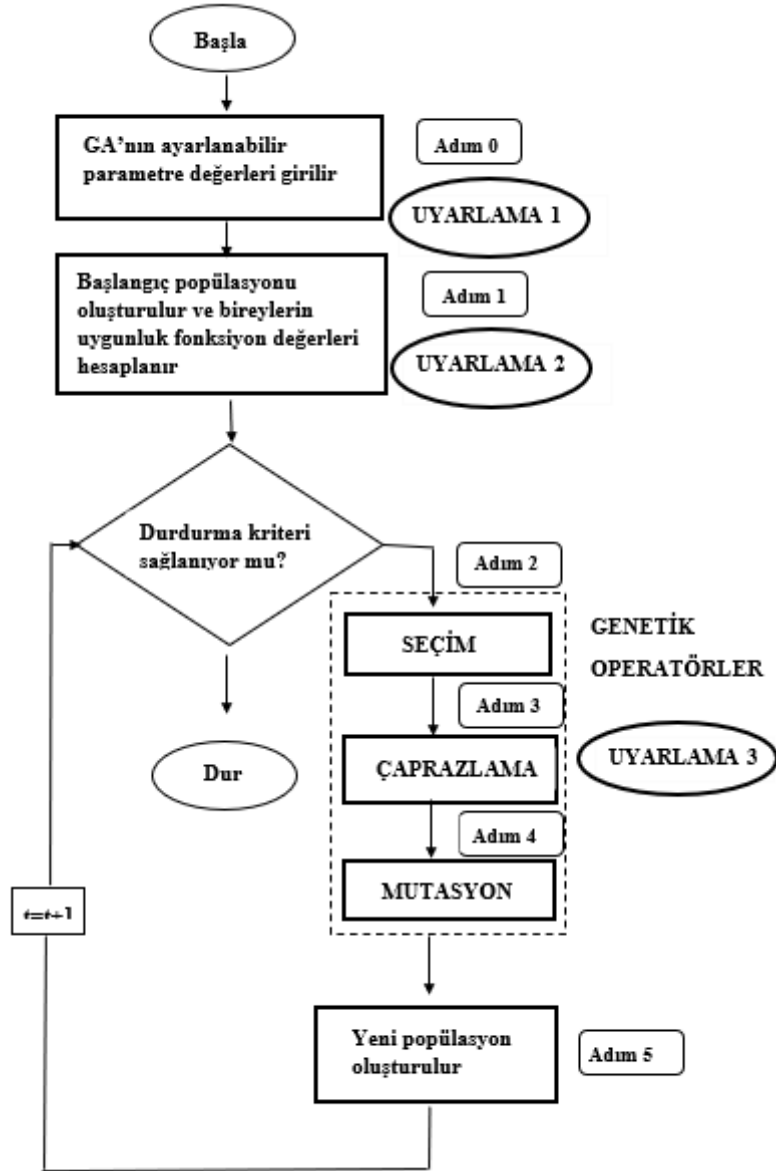
$s'_i = \frac{s_i + s_{i+1}}{2}$ hesaplanarak mutasyon sonrasında elde edilen yeni çözüm

$$c = \begin{cases} s_i & , \quad y + \delta_q (ub - lb) \leq s'_i \\ s_{i+1} & , \quad d.y. \end{cases} \quad (5.7)$$

eşitliği ile elde edilir (Karakoç ve Türkşen 2023).

5.2 Tek Amaçlı Karma Optimizasyon Problemleri için Uyarlanmış Genetik Algoritma

Karma verili tek amaçlı optimizasyon probleminin GA ile çözümü GA'nın karar değişkenlerinin gösterimi, başlangıç popülasyonu oluşturma ve genetik operatörlerin uygulanması aşamalarında uyarlamalar yapılarak oluşturulur.



Şekil 5.1 Uyarlanmış GA'nın akış şeması

Akış şeması şekil 5.1'de verilen uyarlanmış GA'nın algoritmik adımları detaylı biçimde aşağıda ifade edilmiştir.

Adım 0 (UYARLAMA 1): Karar değişkeninin kesikli (tam sayı ve/veya rasyonel sayı) değer alması durumunda karar değişkeni gösteriminde uyarlamalar yapılarak uyarlanmış karar değişkenleri kullanılır.

Adım 1 (UYARLAMA 2): Kesikli değişkenler için başlangıç popülasyonu oluşturulurken değişken gösteriminde dizilere yapılan eşlemeye göre uyarlama yapılır. Uyarlanmış başlangıç popülasyonu oluşturulur.

Adım 2: Ebeveyn seçimi GA'daki gibi yapılır.

Adım 3 (UYARLAMA 3): Eşleşme havuzundaki ebeveynlerle Laplace Çaprazlaması yapılarak yavru bireyler üretilir.

Laplace Çaprazlaması: Laplace Çaprazlama operatörü Laplace dağılımını kullanır. $y^1 = (y_1^1, y_2^1, \dots, y_n^1)$ ve $y^2 = (y_1^2, y_2^2, \dots, y_n^2)$ iki yavru, $x^1 = (x_1^1, x_2^1, \dots, x_n^1)$ ve $x^2 = (x_1^2, x_2^2, \dots, x_n^2)$ iki ebeveynden üretilir İlk olarak $[0,1]$ aralığından rastgele u sayısı üretilir. Laplace dağılımı kullanılarak bir β parametresi bulunur.

$$\beta = \begin{cases} a - b \log(u) & , \quad u \leq 1/2 \\ a + b \log(u) & , \quad u > 1/2 \end{cases} \quad (5.8)$$

Burada, a , konum parametresi ve $b(>0)$, pozitif bir gerçel sayı olup ölçeklendirme parametresidir. Girdi değişkenlerinin tam sayılı olma kısıtlaması varsa $b = b_{int}$. Aksi halde, $b = b_{real}$ alınır. Daha küçük b değerlerinde yavruların ebeveynlere daha yakın üretilmesi muhtemeldir ve daha büyük b değerlerinde de yavruların ebeveynlerden uzakta üretilmesi beklenir. Yavru çözümler

$$\begin{aligned} y_i^1 &= x_i^1 + \beta |x_i^1 - x_i^2| \\ y_i^2 &= x_i^2 + \beta |x_i^1 - x_i^2| \end{aligned} \quad (5.9)$$

biçiminde hesaplanır (Deep ve Thakur 2007a, Deep vd. 2009).

Adım 4 (UYARLAMA 3): Laplace Çaprazlamasıyla üretilen bazı yavru bireylere Güç Mutasyonu uygulanır.

Güç Mutasyonu: Güç Mutasyonu güç dağılımına dayanır. Güç dağılımı

$$f(x) = p x^{p-1}, 0 \leq x \leq 1 \quad (5.10)$$

olmak üzere burada p , dağılım indeksidir. Güç Mutasyonu ebeveyn çözümün $\bar{x}$ yakınında bir x çözümü oluşturmak için kullanılır. s_1 , $[0,1]$ aralığından üretilen rastgele bir sayı olmak üzere güç dağılımını izleyen S rastgele sayısı oluşturulur, $s = (s_1)^{\eta_m}$. Burada η_m mutasyon indeksidir. Mutasyonun gücü mutasyon indeksi tarafından yönetilir. Küçük p değerleri için çözümde daha az bozulma beklenirken büyük p değerleri için daha fazla çeşitlilik elde edilir. Tam sayılı karar değişkenleri için η_m değeri $\eta_{m(int)}$, sürekli karar değişkenleri için $\eta_{m(real)}$ olmak üzere mutasyona uğrayan yavru

$$x = \begin{cases} \bar{x} - s(\bar{x} - lb) & , t < r \\ \bar{x} + s(ub - \bar{x}) & , t \geq r \end{cases} \quad (5.11)$$

biçiminde hesaplanır. Burada $t = \frac{\bar{x} - lb}{ub - \bar{x}}$ ve r , $[0,1]$ arasında düzgün dağılmış bir rastgele sayıdır (Deep ve Thakur 2007b, Deep vd. 2009).

Adım 5: Yeni popülasyonu oluşturacak olan bireyler GA'daki gibi seçilir. Algoritma sonlandırma kriteri GA'daki gibidir.

Optimal değeri elde edilmek istenilen yanıt fonksiyonu, amaç fonksiyonu olarak değerlendirildiğinde elde edilen problem MATLAB R2024a programında 'ga' fonksiyonuyla çözülebilir. Karma optimizasyon probleminin çözümünde çaprazlama operatörü için 'Laplace' ve mutasyon operatörü için 'Güç' seçilmelidir (Deep vd. 2009). Algoritma her bir kesikli değişken değerine atanan sıra numarasıyla optimizasyonu gerçekleştirerek optimal girdi değişken değerleri elde edilir.

5.3 Çok Amaçlı Karma Optimizasyon için Uyarlanmış Genetik Algoritma

Bu tez çalışmasında, karma optimizasyon problemlerinin çözümünde NSGA-II ve NSGA-III yöntemlerinin bazı adımlarında uyarlamalar yapılarak karma optimizasyon problemlerinin çözümü için MDNSGA-II ve MDNSGA-III olarak adlandırılan yöntemler önerilmiştir. Aşağıda, önerilen MDNSGA-II ve MDNSGA-III yöntemlerinin algoritmik adımlarına ayrıntılı biçimde yer verilmiştir.

5.3.1 MDNSGA-II yöntemi

NSGA-II Algoritması'nda yer verilen karar değişkenlerinin gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin uygulanması işlemleri için sırasıyla

- *Adım 0*
- *Adım 1*
- *Adım 3*

adımlarında yapılan uyarlamalarla MDNSGA-II yönteminin algoritması tanımlanmıştır. Şekil 5.2'de akış şeması verilen MDNSGA-II yönteminin algoritmik adımları aşağıdaki gibidir.

MDNSGA-II Algoritması:

Adım 0 (UYARLAMA 1): *NSGA-II Algoritması Adım 0*'daki gibi uygulanır. Uyarlanmış karar değişkenleri tanımlanır.

Adım 1 (UYARLAMA 2): Uyarlanmış başlangıç popülasyonu oluşturulur. Uyarlanmış başlangıç popülasyonu için uygunluk fonksiyon değerleri hesaplanır.

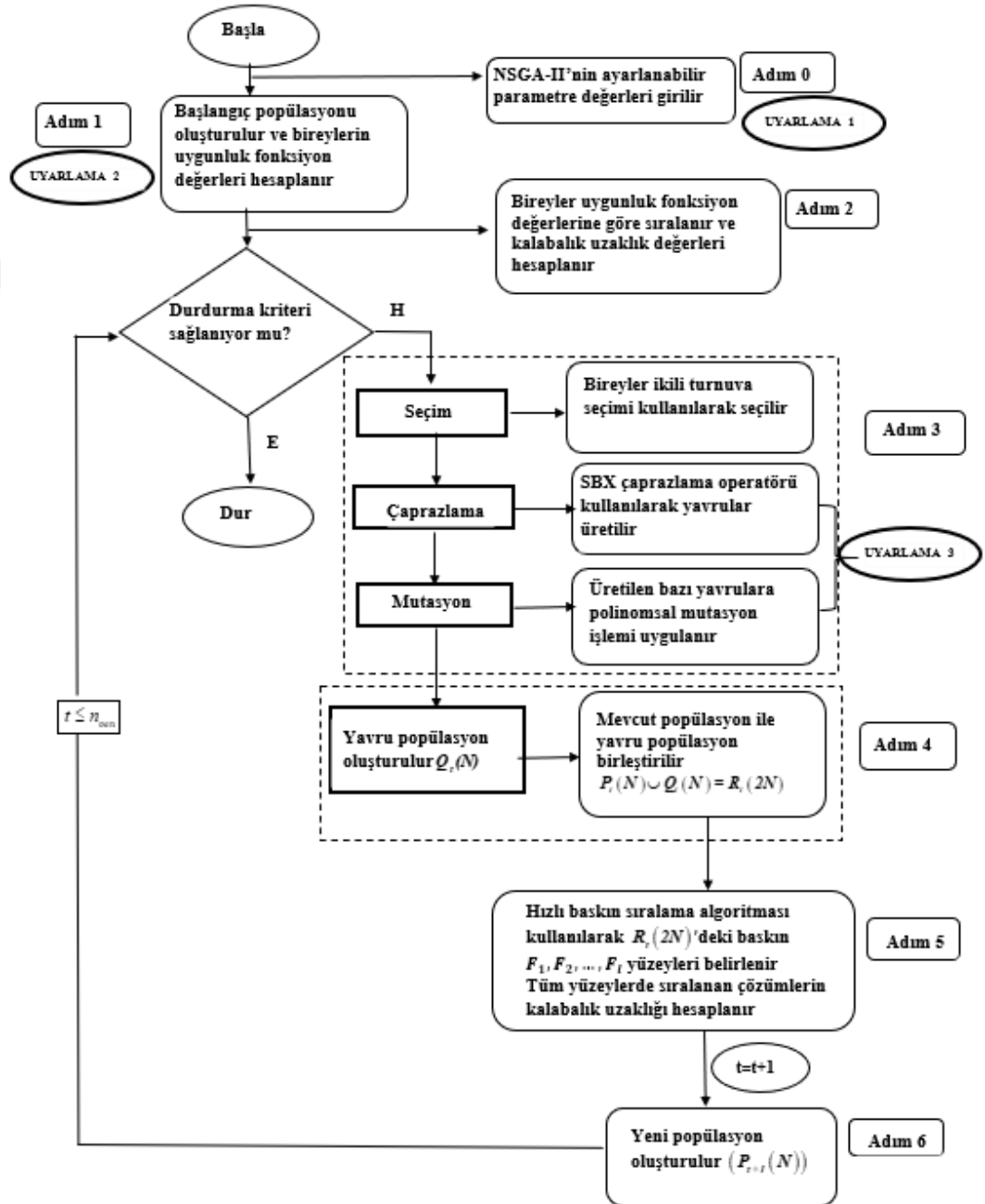
Adım 2: *NSGA-II Algoritması Adım 2*'deki gibi uygulanır.

Adım 3 (UYARLAMA 3): *Adım 0*'da belirtilen ayarlanabilir parametre değerleri kullanılarak uyarlanmış genetik operatörler uygulanır. *NSGA-II Algoritması Adım 3*'teki işleyiş uyarlanmış genetik operatörlerle takip edilir.

Adım 4: NSGA-II Algoritması Adım 4'teki gibi uygulanır.

Adım 5: NSGA-II Algoritması Adım 5'teki gibi uygulanır.

Adım 6: NSGA-II Algoritması Adım 6'daki gibi uygulanır.



Şekil 5.2 MDNSGA-II'nin akış şeması

5.3.2 MDNSGA-III yöntemi

NSGA-III Algoritması’nda yer verilen karar değişkenlerinin gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin uygulanması işlemleri için sırasıyla

- *Adım 0*
- *Adım 1*
- *Adım 3*

adımlarında yapılan uyarlamalarla MDNSGA-III yönteminin algoritması tanımlanmıştır. Şekil 5.3’te akış şeması verilen MDNSGA-III yönteminin algoritmik adımları aşağıdaki gibidir.

MDNSGA-III Algoritması:

Adım 0 (UYARLAMA 1): *NSGA-III Algoritması Adım 0*’daki gibi uygulanır. Uyarlanmış karar değişkenleri tanımlanır.

Adım 1 (UYARLAMA 2): Uyarlanmış başlangıç popülasyonu ve düzgün dağılım ile H sayıda referans noktaları oluşturulur. Uyarlanmış başlangıç popülasyonu için uygunluk fonksiyon değerleri hesaplanır.

Adım 2: *NSGA-III Algoritması Adım 2*’deki gibi uygulanır.

Adım 3 (UYARLAMA 3): *Adım 0*’da belirtilen ayarlanabilir parametre değerleri kullanılarak uyarlanmış genetik operatörler uygulanır. *NSGA-III Algoritması Adım 3*’teki işleyiş uyarlanmış genetik operatörlerle takip edilir.

Adım 4: *NSGA-III Algoritması Adım 4*’teki gibi uygulanır.

Adım 5: *NSGA-III Algoritması Adım 5*’teki gibi uygulanır.

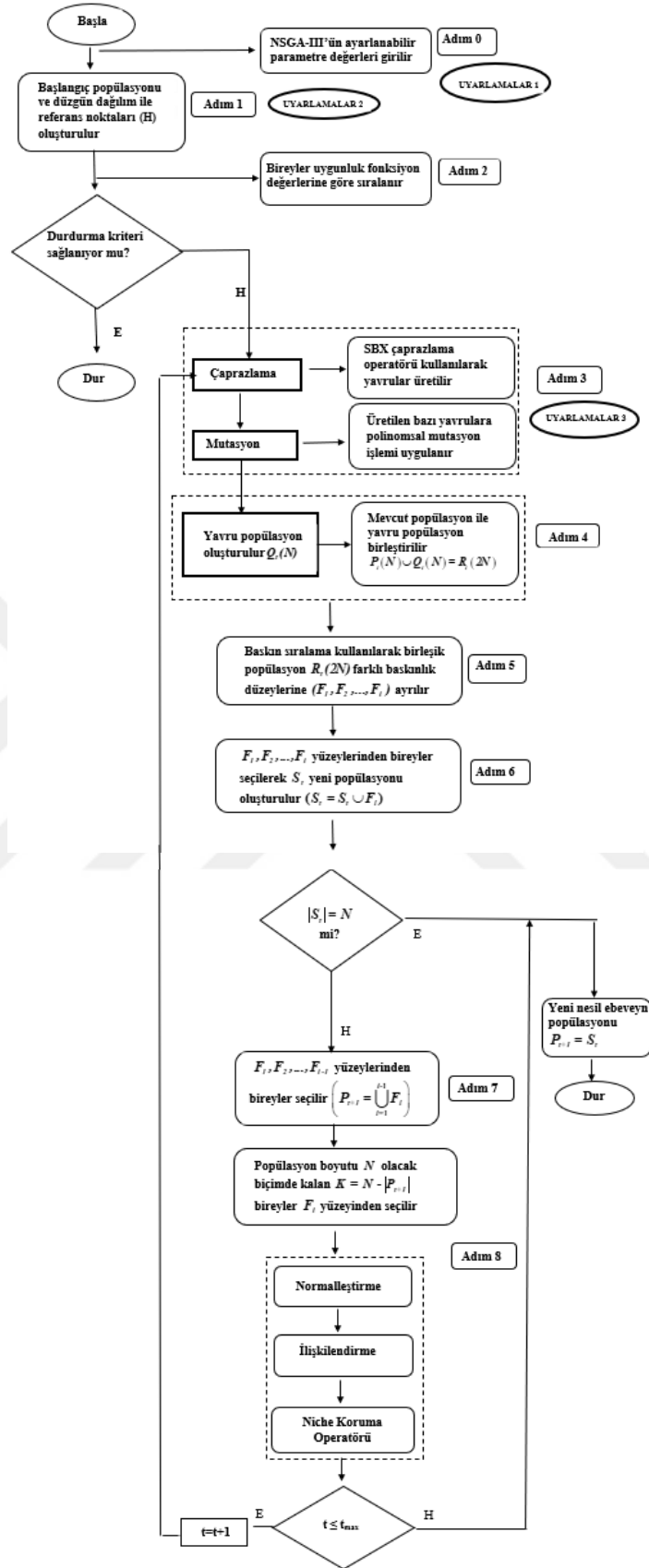
Adım 6: *NSGA-III Algoritması Adım 6*’daki gibi uygulanır.

Adım 7: *NSGA-III Algoritması Adım 7*’deki gibi uygulanır.

Adım 8: *NSGA-III Algoritması Adım 8*'deki gibi uygulanır.

Normalleştirme işlemi, bireylerin amaç fonksiyon değerlerine uygulanır. Popülasyondaki her birey, hiperdüzlemdeki referans çizgileri arasındaki dik mesafeye göre referans noktalarıyla ilişkilendirilir. İlişkilendirme işlemi yapılırken bireylerin sıra numaraları hafızada tutulur. Böylece, amaç fonksiyon değerleriyle işlem yapılırken hafızada tutulan sıra numaraları sayesinde karar değişken değerleri korunur. Niş sayısı küçük olan referans noktalarından K sayıda bireyler seçilir.





Şekil 5.3 MDNSGA-III'ün akış şeması

6. UYGULAMA

Çalışmanın bu bölümünde, karma veri içeren girdi değişkenleri ile birden fazla yanıt değişkenine sahip veri kümelerindeki yanıtların modellenmesine ve MOO uygulamalarına yer verilmiştir. Veri kümesinde yer alan yanıt değişkenlerinin modellenmesinde R 4.3.3 RStudio bilgisayar programı ve IBM SPSS Statistics 24 bilgisayar paket programları kullanılmıştır. Elde edilen tahmini yanıt modellerinin eşanlı optimizasyonu aşamasında girdi değişkenlerinin optimal değerini elde etmek için tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleri MATLAB R2024a ve Python 3.11 Spyder IDE bilgisayar programlarından yararlanılarak uygulanmıştır.

Tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemlerinin uygulanabilirliğinin gösterilmesi amacıyla UCI Machine Learning Repository veri tabanından iki yanıtlı enerji verimliliği veri kümesi ile literatürde tanımlı gıda alanından çok yanıtlı deneysel bir veri kümesi kullanılmıştır. Enerji verimliliği veri kümesinde girdi değişkenleri kesikli ve sürekli değerler aldığından veri kümesi, karma veri kümesi olarak doğrudan ele alınmıştır. Fakat, gıda veri kümesinde uzman görüşü dikkate alınarak girdi değişkenlerinden biri, kesikli değerli olarak kabul edilmiştir. Buna göre, tez çalışmasında iki farklı karma veri kümesi ile çalışılmıştır. Modelleme aşamasında, enerji verimliliği karma veri kümesi için girdi değişkenleri arasında çoklu bağlantı sorunu olması nedeniyle Ridge, LASSO ve Elastik Net regresyon analizi uygulanarak, gıda karma veri kümesi için ise yanıtlar arasında doğrusal ilişki olması nedeniyle SUR analizi uygulanarak tahmini yanıt fonksiyonları elde edilmiştir. MOO aşamasında ise, tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleri probleme özgü ayarlanabilir parametre değerleri tanımlanarak uygulanmıştır.

6.1 Uygulama 1 (Enerji Verimliliği Karma Veri Kümesi)

Fosil enerji kaynaklarının hızla tükendiği ve enerji ihtiyacının hızla arttığı dünyamızda enerji verimliliği kavramının önemi her geçen gün artmaktadır. Gerek fosil enerji kaynaklarının tükenmesi gerekse çevreye olan duyarlılığın artmasıyla birlikte çevreye kısmen daha az zarar veren alternatif enerji kaynaklarının temel enerji kaynakları olarak

kullanılması dünyada birçok devlet tarafından teşvik edilmektedir. Alternatif enerji kaynaklarının daha yaygın kullanılmasının yanı sıra enerji tüketimini en aza indirmek amacıyla enerji verimliliği uygulamaları da aynı anda birçok ülkede teşvik edilmektedir. Özellikle kapalı alanlarda enerji tüketimi en fazla ısıtma ve soğutma işlemleri için harcanmaktadır. Enerji fiyatları da göz önünde bulundurulduğunda insanlar artık ısı yalıtımlı binaları tercih etmektedir. Günümüzde binaların enerji performansı ile ilgili çok sayıda araştırma yapılmaktadır. Yapılan bu araştırmalar sonucunda yayınlanan raporlar, bina enerji tüketiminin dünya çapında istikrarlı bir biçimde arttığını göstermektedir. Sürekli artan enerji talebini azaltmanın bir yolu, enerji verimliliği için bina tasarımlarında ihtiyaç duyulan ısıtma ve soğutma yüklerini hesaplamaktır. Buna göre, ısıtma ve soğutma yükleri enerji verimliliğinde önemli bir yer tutar.

Bu tez çalışmasında, uygulama amacıyla UCI Machine Learning Repository veri tabanından enerji verimliliği konulu karma veri kümesi alınmıştır. Karma veri kümesi 2 yanıt, 8 girdi değişkenli ve 768 gözlem değerlidir. Çalışmada, ısıtma yükü (Y_1) ve soğutma yükü (Y_2) olarak tanımlanan yanıt değişkenlerinin girdi değişkenleriyle arasındaki fonksiyonel ilişkisi belirlenerek yanıtlar modellenmiştir. Elde edilen tahmini yanıt fonksiyonlarının eş anlamlı minimizasyonu yapılmıştır. Çizelge 6.1’de enerji verimliliği konulu karma veri kümesi örneği yer almaktadır.

Çizelge 6.1 Enerji verimliliği karma veri kümesi (UCI Machine Learning Repository)

No	<i>Girdi Değişkenleri</i>								<i>Yanıt Değişkenleri</i>	
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Y_1	Y_2
1	0.98	514.5	294	110.25	7	2	0	0	15.55	21.33
2	0.98	514.5	294	110.25	7	3	0	0	15.55	21.33
3	0.98	514.5	294	110.25	7	4	0	0	15.55	21.33
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
767	0.62	808.5	367.5	220.5	3.5	4	0.4	5	16.48	16.61
768	0.62	808.5	367.5	220.5	3.5	5	0.4	5	16.64	16.03

Çizelge 6.1’de girdi değişkenleri için görülen X_1 bağıl kompaktlık, X_2 yüzey alanı, X_3 duvar alanı, X_4 çatı alanı, X_5 yapı yüksekliği, X_6 yapının yönelimi, X_7 cam alanı ve

X_8 cam alan dağılımı olarak tanımlanmıştır (UCI Maching Learning Repository). Çizelge 6.2’de, girdi ve yanıt değişkenlerinin türleri ile değişkenlerin aldığı değerlerin tanım kümeleri verilmiştir.

Çizelge 6.2 Enerji verimliliği karma veri kümesinin girdi ve yanıt değişkenleri, değişken türleri ve tanım kümeleri

	<i>Değişkenler</i>	<i>Değişken Türü</i>	<i>Tanım Kümesi</i>
Girdiler	Bağıl Kompaktlık (X_1)	Sürekli	[0.62, 0.98]
	Yüzey alanı (X_2)	Sürekli	[514.5, 808.5]
	Duvar alanı (X_3)	Sürekli	[245, 416.5]
	Çatı alanı (X_4)	Sürekli	[110.25, 220.5]
	Yapı yüksekliği (X_5)	Kesikli (Rasyonel sayı)	{3.5, 7}
	Yapının yönelimi (X_6)	Kesikli (Tam sayılı)	{2, 3, 4, 5}
	Cam alanı (X_7)	Kesikli (Rasyonel sayı)	{0, 0.1, 0.25, 0.4}
	Cam alan dağılımı (X_8)	Kesikli (Tam sayılı)	{0, 1, 2, 3, 4, 5}
Yanıtlar	Isıtma Yüğü (Y_1)	Sürekli	$Y_1 \geq 0$
	Soğutma Yüğü (Y_2)	Sürekli	$Y_2 \geq 0$

Çizelge 6.2’den, girdi ve yanıt değişkenlerinin değişken türlerinin birbirinden farklılık gösterdiği açıkça görülmektedir. X_1 , X_2 , X_3 ve X_4 değişkenleri sürekli değer alırken, X_5 ve X_7 değişkenlerinin rasyonel sayılı kesikli değer aldığı, X_6 ve X_8 değişkenlerinin ise tam sayılı kesikli değer aldığı çizelge 6.2’de belirtilmiştir. Çizelge 6.2’de sürekli ve kesikli değer alan karar değişkenlerinin tanım kümeleri açıkça görülmektedir. Buna göre, sürekli değerler alan X_1 , X_2 , X_3 ve X_4 karar değişkenlerinin aldığı değer kümesinin sırasıyla, $X_1 \in [0.62, 0.98]$, $X_2 \in [514.5, 808.5]$, $X_3 \in [245, 416.5]$ ve $X_4 \in [110.25, 220.5]$ olduğu söylenir. Yapı yüksekliği değişkeni için $X_5 \in \{3.5, 7\}$ ’ dir. Binalarda, taban alanının yüzdesi olarak ifade edilen %10, %25 ve %40 olmak üzere üç tip cam alanı (X_7) kullanılmıştır. Bazı binalarda ise hiç cam alanı (%0) bulunmamaktadır. Buna göre, X_7 değişkeninin alabileceği değerler $X_7 \in \{0, 0.1, 0.25, 0.4\}$ olarak verilmiştir. X_6 değişkeni yapının kuzey, güney, doğu ve

batı cephelerinden hangisinde yer aldığını belirtmektedir. Bu cephe kategorilerini belirtmek amacıyla tam sayı değerleri kullanılarak X_6 değişkeninin tanım kümesi $X_6 \in \{2,3,4,5\}$ olarak ifade edilmiştir. Her bir cam alanı için tek tip, kuzey, doğu, güney ve batı olmak üzere beş farklı cam alan dağılımı (X_8) vardır. Ayrıca, cam alanı bulunmadığında ($X_7 = 0$) cam alanının dağılımı olmayacağından cam alan dağılımı değeri 0 (sıfır) olur, $X_8 = 0$. Buna göre X_8 değişkeninin aldığı değerler kümesi $X_8 \in \{0,1,2,3,4,5\}$ 'tir.

Çizelge 6.3'te enerji verimliliği karma veri kümesinin Y_1 ve Y_2 yanıt değişkenleri için hesaplanan betimsel istatistikleri verilmiştir.

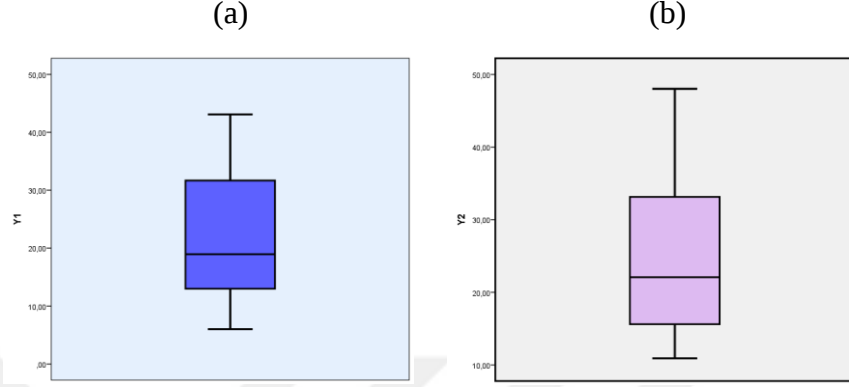
Çizelge 6.3 Y_1 ve Y_2 yanıt değişkenleri için betimsel istatistikler

<i>Betimsel İstatistikler</i>	Y_1	Y_2
<i>Ortalama</i>	22.3	24.587
<i>Standart sapma</i>	10.09	9.513
<i>1.Çeyreklik</i>	12.977	15.6
<i>2.Çeyreklik (Medyan)</i>	18.95	22.08
<i>3.Çeyreklik</i>	31.682	33.137
<i>Minimum</i>	6.01	10.9
<i>Maksimum</i>	43.1	48.03
<i>Range</i>	37.09	37.13
<i>Çarpıklık</i>	0.36	0.396

Çizelge 6.3'ten sürekli değer alan Y_1 ve Y_2 yanıtlarının ortalama ve standart sapma değerlerinin yaklaşık olarak benzer olduğu görülmektedir. Her iki yanıtla ilişkin tanım aralığı uzunluğu (Range) da benzerdir. Çarpıklık katsayısı değerlerine bakıldığında Y_1 ve Y_2 yanıtlarının dağılımının sağa çarpık olduğu söylenebilir. Bu durum, karma veri kümesi için ısıtma ve soğutma yüklerinin düşük değerler etrafında yoğunlaştığını gösterir.

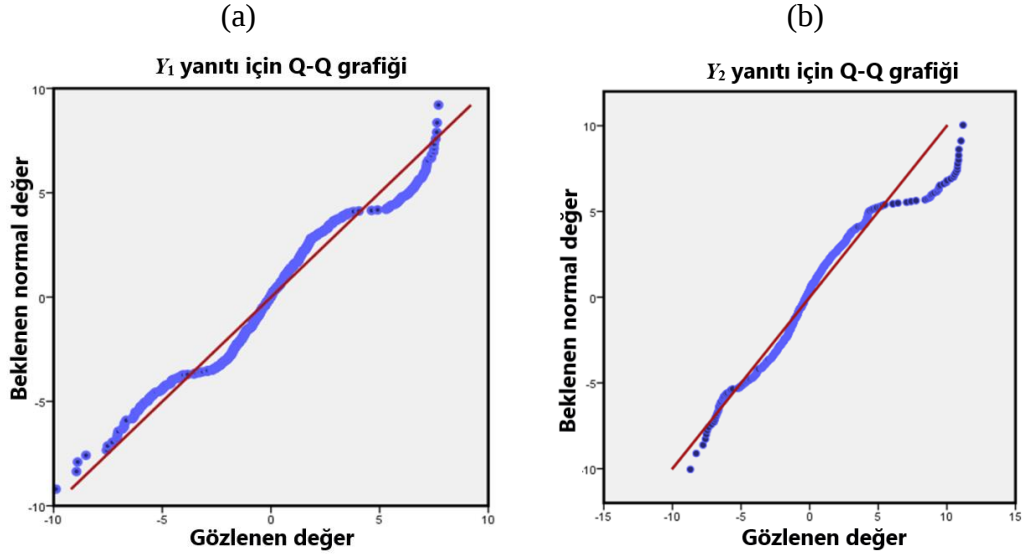
Şekil 6.1'de, Y_1 ve Y_2 yanıtları için kutu grafikleri yer almaktadır. Şekil 6.1 (a)-(b)'deki, kutu grafiklerinden de Y_1 ve Y_2 yanıt değişkenlerinin dağılımının sağa çarpık olduğu

açıkça görülmektedir. Buna göre, Y_1 ve Y_2 yanıt değişkenlerinin simetrik dağılımlı olan Normal dağılıma uyduğu söylenemez. Ayrıca ısıtma ve soğutma yüklerinde aykırı değerler olmadığı söylenir.



Şekil 6.1 (a) Y_1 ve (b) Y_2 yanıt değişkenleri için kutu grafikleri

Y_1 ve Y_2 yanıt değişkenleri için Q-Q grafikleri çizdirilerek yanıtların normal dağılım gösterip göstermediğine bakılmıştır. Şekil 6.2’de görülen Q-Q grafiklerinden Y_1 ve Y_2 yanıtlarına ilişkin gözlem değerlerinin Normal dağılım göstermediği görülmektedir.

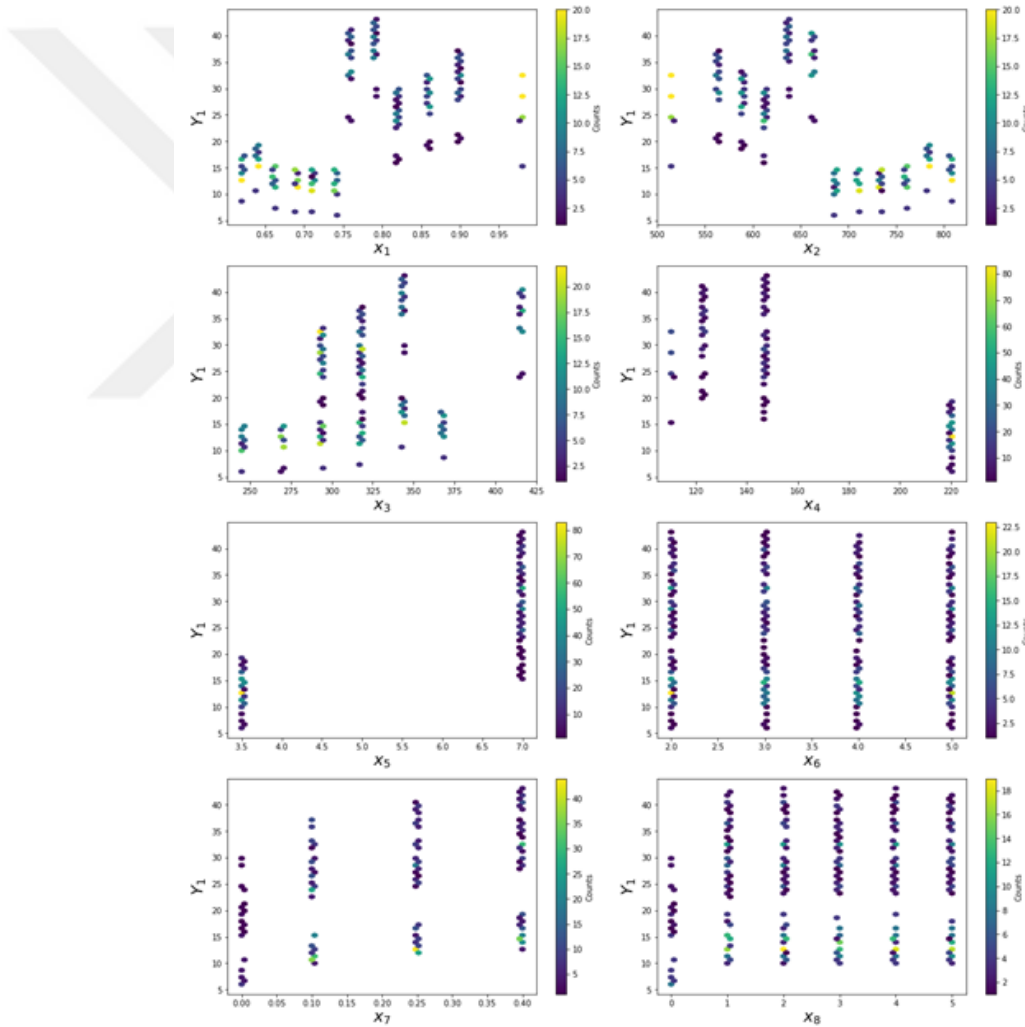


Şekil 6.2 (a) Y_1 ve (b) Y_2 yanıt değişkenleri için Q-Q grafikleri

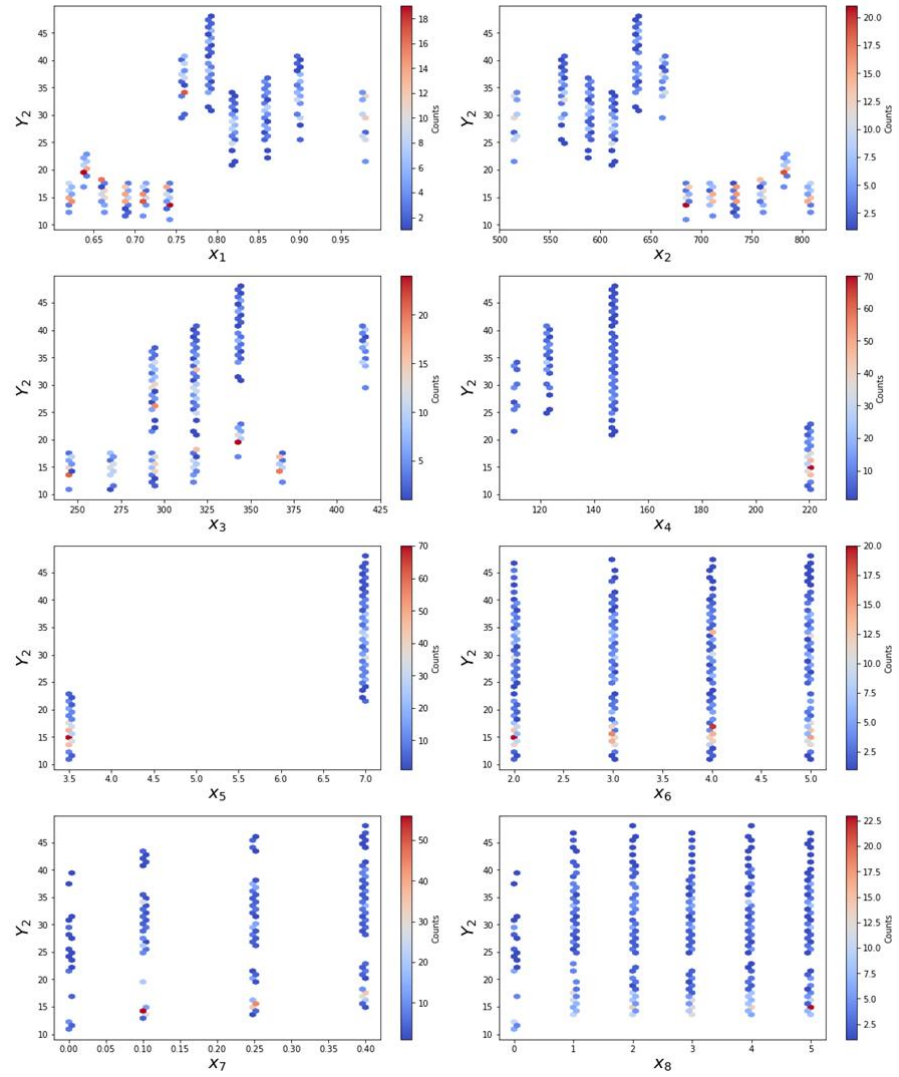
İstatistiksel hipotez testleri uygulandığında, Y_1 ve Y_2 yanıt değişkenlerine uygulanan Kolmogrov-Smirnov testi sonucuna göre de %95 güven düzeyinde Y_1 ve Y_2 yanıtlarına

ilişkin gözlem değerlerinin Normal dağılım göstermediği sonucuna ulaşılmıştır, p -değeri <0.05 .

Y_1 ve Y_2 yanıtlarıyla girdi değişkenleri $\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8\}$ arasındaki ilişkiyi gösteren saçılım grafikleri şekil 6.3-6.4'te verilmiştir. Bu saçılım grafiklerinden girdi ile yanıt değişkenleri arasında ilişki olduğu görülmektedir. Fakat, ilişkinin doğrusal ya da doğrusal olmadığı açık ve net değildir. Bu nedenle, Y_1 ve Y_2 yanıtlarının Normal dağılımlı olmadığı göz önüne alınarak parametrik olmayan korelasyon katsayılarından biri olan Spearman korelasyon katsayıları hesaplanmıştır.



Şekil 6.3 Y_1 yanıtıyla girdi değişkenleri arasındaki ilişkiyi gösteren saçılım grafikleri



Şekil 6.4 Y_2 yanıtıyla girdi değişkenleri arasındaki ilişkiyi gösteren saçılım grafikleri

Girdi ve yanıt değişkenleri arasındaki Spearman sıra korelasyon değerlerini içeren korelasyon matrisi Python 3.11 Spyder programında hesaplanarak ısı haritası elde edilmiştir. Şekil 6.5'te verilen ısı haritasına göre girdi ve yanıt değişkenleri arasında doğrusal ilişki olduğu görülürken, girdi değişkenleri arasında da doğrusal ilişki olduğu görülmektedir. Girdi değişkenleri arasında doğrusal bağlantı olması durumu çoklu bağlantı sorunu olduğunu gösterir. Şekil 6.5'te verilen ısı haritasındaki Spearman sıra korelasyon değerlerinden $X_1 - X_2$, $X_1 - X_4$, $X_1 - X_5$, $X_2 - X_4$, $X_2 - X_5$ ve $X_4 - X_5$ girdi değişkenleri arasında doğrusal bağlantı olduğu görülmektedir.



Şekil 6.5 Girdi ve yanıt değişkenleri arasındaki korelasyon matrisi ısı haritası (Spearman sıra korelasyonu)

Bu tez çalışmasında, çoklu bağlantı sorununu gidermek amacıyla Ridge (R), LASSO (L) ve Elastik Net (EN) regresyon yöntemleri uygulanarak enerji verimliliği karma veri kümesinde Y_1 ve Y_2 yanıtları modellenmiştir. Y_1 ve Y_2 yanıtları için elde edilen tahmini yanıt modelleri sırasıyla $\hat{Y}_1^R$ ve $\hat{Y}_2^R$ olmak üzere

$$\hat{Y}_1^R = -18.8468x_1 - 0.0045x_2 + 0.0345x_3 - 0.0195x_4 + 5.13x_5 - 0.022x_6 + 20.1451x_7 + 0.2188x_8 \quad (6.1)$$

$$\hat{Y}_2^R = -16.5927x_1 + 0.0007x_2 + 0.0218x_3 - 0.0105x_4 + 5.2931x_5 + 0.0669x_6 + 14.7774x_7 + 0.0432x_8 \quad (6.2)$$

$$\hat{Y}_1^L = -16.0196x_1 + 0.0132x_2 + 0.0192x_3 - 0.0614x_4 + 4.8073x_5 - 0.0302x_6 + 19.6476x_7 + 0.2167x_8 \quad (6.3)$$

$$\hat{Y}_2^L = -14.0662x_1 + 0.0125x_2 + 0.0124x_3 - 0.0395x_4 + 5.0073x_5 + 0.0444x_6 + 14.3285x_7 + 0.0421x_8 \quad (6.4)$$

$$\hat{Y}_1^{EN} = -18.9282x_1 + 0.013x_2 + 0.017x_3 - 0.0546x_4 + 5.1354x_5 - 0.0227x_6 + 20.1851x_7 + 0.2178x_8 \quad (6.5)$$

$$\hat{Y}_2^{EN} = -16.8643x_1 + 0.0122x_2 + 0.0103x_3 - 0.0331x_4 + 5.3192x_5 + 0.0672x_6 + 14.8473x_7 + 0.0421x_8 \quad (6.6)$$

dır. Eşitlik (6.1)-(6.6) ile verilen $\hat{Y}_1$ ve $\hat{Y}_2$ modellerinin performansının değerlendirilmesinde belirtme katsayısı R^2 , ortalama hata kareler karekökü (Root Mean Square Error-*RMSE*) ve ortalama mutlak hata (Mean Absolute Error-*MAE*) metrikleri kullanılmıştır. Bu metrikler, sırasıyla,

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}, \quad (6.7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (6.8)$$

ve

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (6.9)$$

biçiminde tanımlıdır. Yanıt değişkenindeki değişimin ne kadarının girdi değişkenleri ile açıklandığını belirten R^2 belirtme katsayısı değerinin yüksek olması, bunun yanı sıra *MAE* ve *RMSE* değerlerinin ise düşük olması durumunda tahmini yanıt modellerinin iyi bir performans gösterdiği düşünülür. Ridge, LASSO ve Elastik Net regresyon yöntemleriyle elde edilen modellerin performans metrik değerleri çizelge 6.4'te verilmiştir. Çizelge 6.4'te görülen performans değerlerinden $\hat{Y}_1$ modeli için Ridge ve $\hat{Y}_2$ modeli için Elastik Net regresyon yöntemlerinin performansının iyi olduğu sonucuna

ulaşılır. Buna göre, $\hat{Y}_1$ modeli için eşitlik (6.1) ve $\hat{Y}_2$ modeli için eşitlik (6.6) tercih edilir. $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modelleri için artık analizi yapılarak model uygunluğu değerlendirilmiştir.

Çizelge 6.4 $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modelleri için elde edilen performans değerleri

<i>Yöntemler</i>	$\hat{Y}_1^R$			$\hat{Y}_2^{EN}$		
	R^2	MAE	RMSE	R^2	MAE	RMSE
Ridge	0.9091	3.0775	2.2729	0.8904	3.1860	2.2457
LASSO	0.9083	3.0912	2.2565	0.8886	3.2127	2.2681
Elastik Net	0.9091	3.0776	2.2737	0.8904	3.1842	2.2444

Artıklar arasında ilişki olup olmadığı Durbin-Watson test istatistiğiyle kontrol edilmiştir. $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modelleri için Durbin-Watson test istatistikleri sırasıyla 0.61 ile 1.0382 olarak hesaplanmıştır. Hesaplanan bu değerler [0,4] aralığında olduğundan artıkların ilişkisiz olduğu söylenir.

Artıkların varyansının sabit olup olmadığını test etmek için White testi uygulanmıştır. Test sonucuna göre, %95 güven düzeyinde $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ model artıklarının varyansının sabit olmadığı sonucuna ulaşılmıştır, p -değeri = 0.0001 < 0.05. Bu istenmeyen bir durumdur. Fakat, regresyon katsayısı tahminlerinde yanlışlığa neden olmaz (Sen ve Srivastava 1990). Bu tez çalışmasında, elde edilen $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ tahmini yanıt fonksiyonlarının eş anlı optimizasyonunu hedefleyen çok yanıtlı optimizasyon probleminde, yanıt fonksiyonları sırasıyla f_1 ve f_2 amaç fonksiyonları olarak ele alındığında, problem

$$\begin{aligned} \min f_1 &= \hat{Y}_1^R(\mathbf{x}) \\ \min f_2 &= \hat{Y}_2^{EN}(\mathbf{x}) \\ \mathbf{x} &\in S \end{aligned} \quad (6.10)$$

biçiminde tanımlı MOO problemine dönüşür. (6.10) ifadesinde belirtilen MOO probleminin çözümü için tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleri uygulanmıştır. Yöntemlerin kullanılmasında literatür bilgisi ve uzman görüşü kullanılarak ayarlanabilir parametre değerleri belirlenmiştir. MDNSGA-II ve MDNSGA-

III yöntemlerine ilişkin ayarlanabilir parametrelerin aldığı değerler sırasıyla, çizelge 6.5 ve çizelge 6.6’da verilmiştir.

Çizelge 6.5 MDNSGA-II’nin ayarlanabilir parametre değerleri

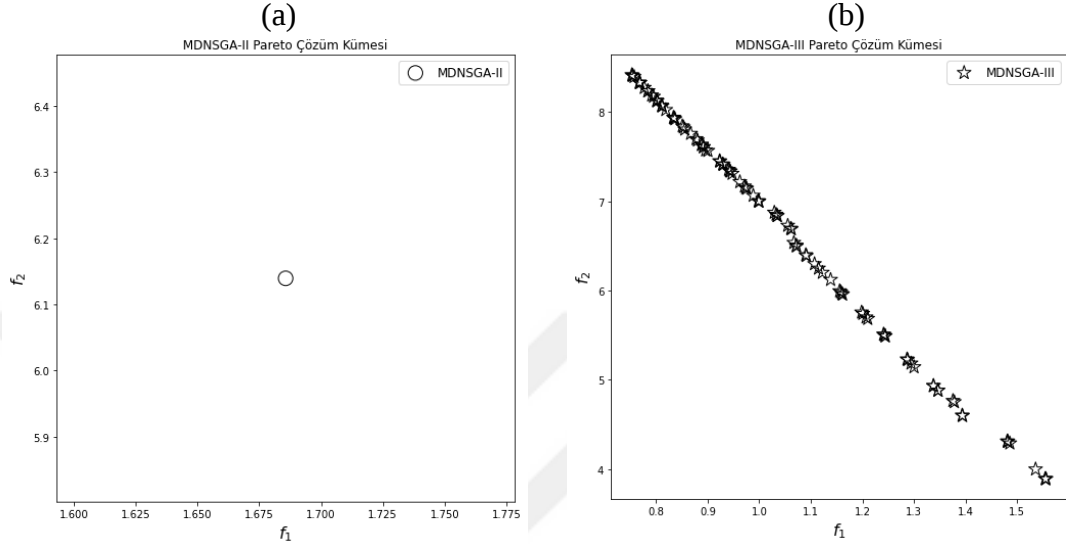
<i>Ayarlanabilir parametreler</i>	<i>Parametre değerleri</i>
Girdi değişken sayısı (p)	8
Popülasyon büyüklüğü (N)	100
Yineleme sayısı (n_{gen})	200
Çaprazlama olasılığı (P_c)	0.90
Çaprazlama indeksi (η_c)	20
Mutasyon olasılığı (P_m)	1/8
Mutasyon indeksi (η_m)	20
Çaprazlama operatörü	SBX
Mutasyon operatörü	Polinomsal
Seçim operatörü	Turnuva

Belirlenen ayarlanabilir parametre değerleriyle, MDNSGA-II yöntemini MATLAB R2024a’da ve MDNSGA-III yöntemini Python 3.11 Spyder’da uygulayarak MOO problemi optimize edilmiştir. MOO sonucunda elde edilen Pareto çözüm kümesi şekil 6.6 (a)-(b)’de görülmektedir.

Çizelge 6.6 MDNSGA-III’ün ayarlanabilir parametre değerleri

<i>Ayarlanabilir parametreler</i>	<i>Parametre değerleri</i>
Girdi değişken sayısı (p)	8
Popülasyon büyüklüğü (N)	100
Yineleme sayısı (n_{gen})	200
Referans bölüm sayısı (n_{div})	24
Çaprazlama olasılığı (P_c)	0.90
Çaprazlama indeksi (η_c)	20
Mutasyon olasılığı (P_m)	1/8
Mutasyon indeksi (η_m)	20
Çaprazlama operatörü	SBX
Mutasyon operatörü	Polinomsal

Şekil 6.6 (a)'dan, MDNSGA-II ile elde edilen Pareto çözüm kümesinin tek bir nokta içerdiği gözlenirken Şekil 6.6 (b)'den, MDNSGA-III ile elde edilen Pareto çözüm kümesinin çok sayıda çözüm içerdiği görülmektedir. Buna göre, MDNSGA-III ile Pareto çözüm kümesinde çözüm çeşitliliğinin sağlandığı sonucuna ulaşılır.

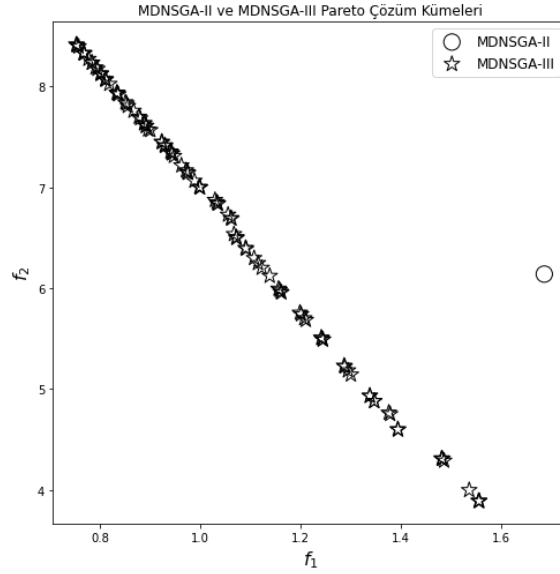


Şekil 6.6 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için elde edilen Pareto çözüm kümeleri

Şekil 6.7'de MDNSGA-II ve MDNSGA-III yöntemleriyle elde edilen Pareto çözüm kümeleri birlikte verilmiştir. Şekil 6.7'den, MDNSGA-III ile çözüm çeşitliliğinin sağlanmasının yanı sıra MDNSGA-II'ye göre baskın çözümler elde edildiği görülmüştür. MDNSGA-II'nin Pareto çözüm kümesi Çizelge 6.7'de verilmiştir. Çizelge 6.7'de, girdi değişkenlerinin $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ ve amaç fonksiyonu vektörünün $f = [f_1 \ f_2]$ aldığı değerler görülmektedir. MDNSGA-III ile elde edilen Pareto çözüm kümesine EK 1'de yer verilmiştir.

Çizelge 6.7 Enerji verimliliği karma verisinde $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için MDNSGA-II ile elde edilen Pareto çözüm kümesi

No	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	$[f_1 \ f_2]$
1	0.98	645.74	301.38	220.5	3.5	3	0	4	[1.68 6.14]



Şekil 6.7 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ modeli için elde edilen Pareto çözüm kümelerinin karşılaştırılması

Elde edilen Pareto çözüm kümesinin uygulanan MOO yöntemine bağlı olduğu açıktır. Bununla birlikte farklı amaç fonksiyonları için de Pareto çözüm kümesi değişir. Enerji verimliliği karma veri kümesindeki yanıtların modellenmesi amacıyla karma veri yapısına uygun olarak literatürde farklı bir modelleme çalışması yapılmıştır. Tsanas ve Xifara (2012), verilerin Normal dağılım göstermediği, aykırı değerlerin varlığı durumunda İteratifli Yeniden Ağırlıklandırılmış En Küçük Kareler (Iteratively Reweighted Least Squares-IRLS) yöntemini kullanmanın model için uygun olduğunu belirtmişlerdir. Yapmış oldukları çalışmada enerji verimliliği karma veri kümesi için IRLS yöntemini kullanarak Y_1 ve Y_2 yanıtları için elde ettikleri tahmini yanıt fonksiyonları sırasıyla

$$\hat{Y}_1^{IRLS} = -4.75x_1 - 0.03x_2 + 0.07x_3 - 3.44x_5 - 0.01x_6 + 18.13x_7 + 0.09x_8 \quad (6.11)$$

ve

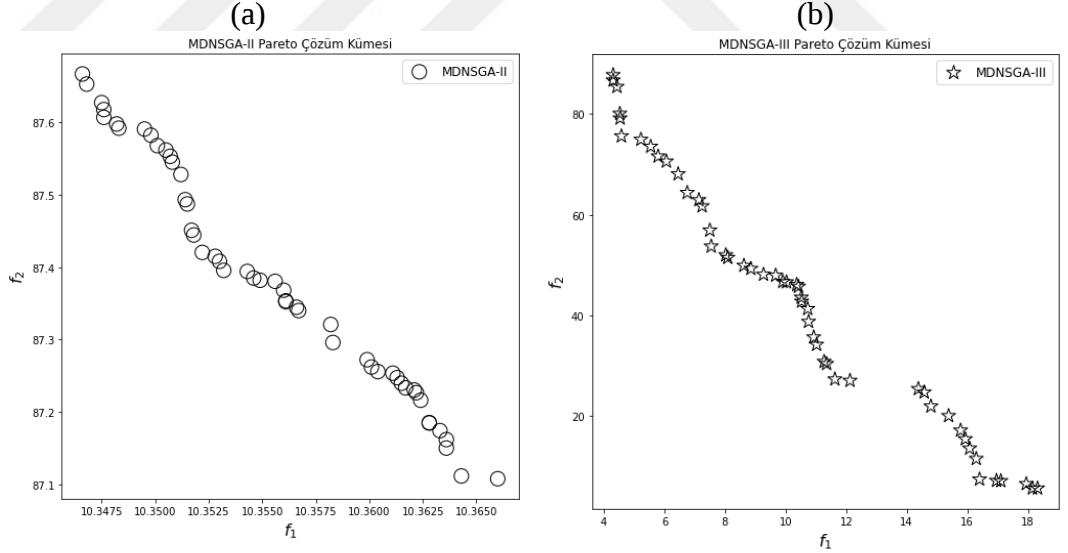
$$\hat{Y}_2^{IRLS} = -9.02x_1 - 0.01x_2 + 0.04x_3 - 4.30x_5 - 0.12x_6 + 14.49x_7 + 0.03x_8 \quad (6.12)$$

dir. IRLS ile regresyon katsayılarındaki ağırlıkları ayarlayarak regresyon eğrisi oluşturulurken aykırı değerlerin etkisini azaltmışlardır.

$\hat{Y}_1^{\text{IRLS}}$ için $MAE = 2.14$, $RMSE = 2.21$ ve $\hat{Y}_2^{\text{IRLS}}$ için $MAE = 3.14$, $RMSE = 3.38$ olarak hesaplanmıştır. Eşitlik (6.11)-(6.12)'de belirtilen $\hat{Y}_1^{\text{IRLS}}$ ve $\hat{Y}_2^{\text{IRLS}}$ modelleri amaç fonksiyonları olarak ele alındığında

$$\begin{aligned} \min f_1 &= \hat{Y}_1^{\text{IRLS}}(\mathbf{x}) \\ \min f_2 &= \hat{Y}_2^{\text{IRLS}}(\mathbf{x}) \\ \mathbf{x} &\in S \end{aligned} \quad (6.13)$$

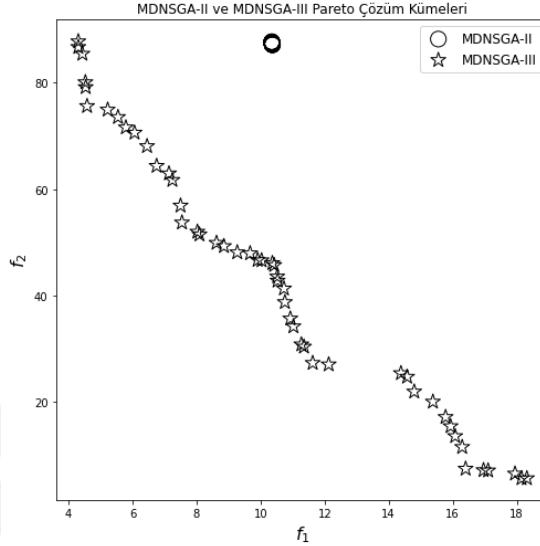
biçimde tanımlı MOO problemine dönüşür. MDNSGA-II ve MDNSGA-III yöntemleri (6.13) ifadesine çizelge (6.5)- (6.6)'da tanımlı ayarlanabilir parametre değerlerinde bazı değişiklikler yapılarak uygulanmıştır. Buna göre, burada MOO problemine uygun olarak $N = 50$, $n_{\text{gen}} = 100$ ve $n_{\text{div}} = 12$ alınmıştır. MDNSGA-II ve MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümeleri şekil 6.8 (a)-(b)'de görülmektedir. Şekil 6.8 (a)-(b)'den MDNSGA-II ve MDNSGA-III yöntemleriyle Pareto çözüm kümelerinde çözüm çeşitliliğinin sağlandığı sonucuna ulaşılmıştır.



Şekil 6.8 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümeleri

Şekil 6.9'da, MDNSGA-II ve MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümelerinin karşılaştırılması verilmiştir. Buna göre, MDNSGA-II ile elde edilen Pareto çözüm kümesindeki amaç fonksiyon değerlerinin daha dar bir aralıkta

çözüm çeşitliliğini sağladığı görülürken MDNSGA-III ile sağlanan çözüm çeşitliliğinin amaç fonksiyonlarının daha geniş değerli biçimde sağlandığı görülmüştür. IRLS modeli için MDNSGA-II ve MDNSGA-III yöntemleriyle elde edilen Pareto çözüm kümelerine sırasıyla, EK 2 ve EK 3'te yer verilmiştir.



Şekil 6.9 MDNSGA-II ve MDNSGA-III yöntemleriyle IRLS modeli için elde edilen Pareto çözüm kümelerinin karşılaştırılması

6.2 Uygulama 2 (Gıda Karma Veri Kümesi)

Sistein jel dokusu ve bir tuz çeşidi olan kalsiyum klorür, peynir altı suyunda yer alan maddelerdir. Bu maddelerin peynir altı suyu konsantresinin yapısal özellikleri ve su tutma özelliği üzerine etkisini incelemek amacıyla yapılan deneysel çalışmada, peynir altı suyu konsantresinin yapısal özellikleri, sertlik, yapışkanlık ve esneklik ile su tutma özelliği sıkıştırılabilir su olarak tanımlanmıştır. Khuri ve Cornell (1996) ile Schmidt vd.'nin (1979) çalışmalarında kullandıkları gıda alanından çok yanıtlı deneysel veri kümesi, 2 girdi ve 4 yanıt değişkeni içermektedir. Sistein jel dokusu (X_1) ve kalsiyum klorür (X_2) girdi değişkenleri ile sertlik (Y_1), yapışkanlık (Y_2), esneklik (Y_3) ve sıkıştırılabilir su (Y_4) yanıt değişkenleri olarak belirlenmiştir. Yapılan deneysel çalışmada, Y_1, Y_2 ile Y_3 yanıtlarını maksimum ve Y_4 yanıtını minimum yapan X_1 ve X_2 girdi değişken değerlerinin belirlenmesi amaçlanmıştır. Deneysel veri kümesinde yer alan X_1 ve X_2 girdi değişkenlerinin gerçek düzeyleri ve kodlanmış düzeyleri çizelge 6.8'de verilmiştir.

Buradan X_1 ve X_2 için deneyde kullanılan gerçek düzeyleri sırasıyla, X_1^{orj} ve X_2^{orj} , X_1 ve X_2 için kodlanmış düzeyleri sırasıyla, X_1^{kod} ve X_2^{kod} olarak gösterilmektedir.

Çizelge 6.8 Girdi değişkenlerinin gerçek ve kodlanmış düzeyleri

X_1^{orj}	2.6	8.0	21.0	34.0	39.4
X_2^{orj}	2.5	6.5	16.2	25.9	29.9
X_1^{kod} ve X_2^{kod}	-1.414	-1	0	1	1.414

Çizelge 6.8’de X_1^{orj} için 21.0 ve X_2^{orj} için 16.2 düzeyleri merkez olmak üzere, girdi değişkenlerinin kod düzeyleri, $X_1^{kod} = (X_1^{orj} - 21) / ((34.0 - 8.0) / 2)$ ve $X_2^{kod} = (X_2^{orj} - 16.2) / ((25.9 - 6.5) / 2)$ eşitliklerinden yararlanılarak elde edilmiştir.

Çizelge 6.9’da, deneysel çalışma sonucunda elde edilen kodlanmış girdi değişken düzeyleri ile gözlenen yanıt değişkenlerine ait değerler yer almaktadır. Çizelge 6.9’daki veri kümesi Merkezi Bileşik Tasarımı (Central Composite Design-CCD) kullanılarak oluşturulmuştur.

Çizelge 6.9 Deneysel çalışma sonucu elde edilen kodlanmış girdi değişken düzeyleri ve gözlenen yanıt değişken değerleri

No	X_1	X_2	Y_1	Y_2	Y_3	Y_4
1	-1	-1	2.48	0.55	1.95	0.22
2	1	-1	0.91	0.52	1.37	0.67
3	-1	1	0.71	0.67	1.74	0.57
4	1	1	0.41	0.36	1.20	0.69
5	-1.414	0	2.28	0.59	1.75	0.33
6	1.414	0	0.35	0.31	1.13	0.67
7	0	-1.414	2.14	0.54	1.68	0.42
8	0	1.414	0.78	0.51	1.51	0.57
9	0	0	1.50	0.66	1.80	0.44
10	0	0	1.66	0.66	1.79	0.50
11	0	0	1.48	0.66	1.79	0.50
12	0	0	1.41	0.66	1.77	0.43
13	0	0	1.58	0.66	1.73	0.47

Literatürde mevcut olan bu deneysel veri kümesinde, girdi değişkenlerinden Sistein jel dokusu (X_1), rasyonel değerli kesikli değişken olarak kabul edilerek karma veri kümesi oluşturulmuştur.

Gıda karma veri kümesi için girdi değişkenleri, değişken türleri ve tanım kümeleri çizelge 6.10’da verilmiştir. X_1 kesikli değişkeni rasyonel değerler alırken X_2 değişkeni sürekli değerler almaktadır. X_1 ve X_2 değişkeninin aldığı değerler kümesi sırasıyla $\{-1.414, -1, 0, 1, 1.414\}$ ve $[-1.414, 1.414]$ olduğu söylenir.

Çizelge 6.10 Gıda karma verisi için girdi değişkenleri, değişken türü ve tanım kümesi

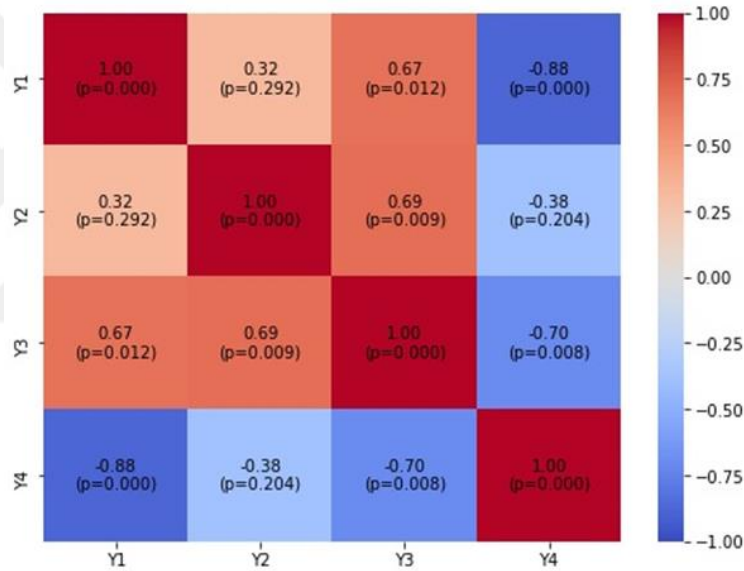
	<i>Girdi Değişkenleri</i>	<i>Değişken Türü</i>	<i>Tanım Kümesi</i>
Girdiler	Sistein jel dokusu (X_1)	Kesikli (Rasyonel sayı)	$\{-1.414, -1, 0, 1, 1.414\}$
	Kalsiyum klorür (X_2)	Sürekli	$[-1.414, 1.414]$
Yanıtlar	Sertlik (Y_1)	Sürekli	$Y_1 > 0$
	Yapışkanlık (Y_2)	Sürekli	$Y_2 > 0$
	Esneklik (Y_3)	Sürekli	$Y_3 > 0$
	Sıkıştırılabilir su (Y_4)	Sürekli	$Y_4 > 0$

Peynir altı suyu proteini konsantresi sistemlerinin yanıt değişkenleri üzerindeki etkilerini ölçmek için oluşturulan gıda karma veri kümesinde yanıtların modellenmesinden önce çok sayıda yanıtlar arasında ilişki olup olmadığına bakılmıştır. Bu amaçla uygun korelasyon katsayısının belirlenebilmesi için öncelikle yanıt değişkenlerinin Normal dağılımlı olup olmadığı incelenmiştir. Yanıt değişkenlerinin Normal dağılım varsayımı için Shapiro-Wilk (küçük örneklemelerde) testi uygulanmıştır. Shapiro-Wilk testi sonucu çizelge 6.11’de özetlenmiştir. Çizelge 6.11’deki p -değerlerine bakıldığında %95 güven düzeyinde Y_1 ve Y_4 yanıtlarının Normal dağılımlı olduğu (p -değeri >0.05), Y_2 ve Y_3 yanıtlarının dağılımının ise Normal dağılım göstermediği söylenir (p -değeri <0.05).

Çizelge 6.11 Gıda karma verisi için yanıt değişkenlerinin Shapiro-Wilk Normallik testi

<i>Yanıt Değişkenleri</i>	<i>Shapiro-Wilk p-değeri</i>
Y_1	0.549
Y_2	0.010
Y_3	0.023
Y_4	0.652

Bazı yanıtların Normal dağılımlı bazılarının ise Normal dağılımlı olmaması nedeniyle yanıtlar arasındaki doğrusal ilişkinin büyüklüğünün belirlenmesi için Pearson ve Spearman korelasyon katsayıları hesaplanmıştır. Normal dağılımlı olmayan yanıt değişkenleri ile tüm yanıt değişkenleri arasındaki ilişkiyi gösteren Spearman korelasyon matrisi ısı haritası şekil 6.10'da görülmektedir. Şekil 6.10'daki korelasyon değerlerine göre $Y_1 - Y_2$ ve $Y_2 - Y_4$ yanıtlarının doğrusal ilişkisiz olduğu görülmektedir. $Y_1 - Y_4$ yanıt değişkenleri için Pearson korelasyon katsayısı -0.932 (p -değeri=0.0001<0.5) olup bu yanıtlar arasında ters yönlü, güçlü ve anlamlı bir ilişki olduğu söylenir. $Y_1 - Y_3$, $Y_2 - Y_3$ ve $Y_3 - Y_4$ yanıtları arasında sırasıyla, 0.67, 0.69 ve -0.70 büyüklüklerinde %95 güven düzeyinde istatistiksel olarak anlamlı bulunan doğrusal ilişki olduğu söylenir.



Şekil 6.10 Yanıt değişkenleri arasındaki korelasyon matrisi ısı haritası

Bu tez çalışmasında, Normallik varsayımını sağladığı ve aralarında anlamlı, güçlü bir ilişki olduğu için $Y_1 - Y_4$ yanıtları, aralarında anlamlı, pozitif yönlü doğrusal bir ilişki olduğu için $Y_2 - Y_3$ yanıtları ve aralarında anlamlı, negatif yönlü doğrusal bir ilişki olduğu için $Y_3 - Y_4$ yanıtları ile çalışılmıştır. Yanıtlar arasındaki ilişki, ikinci dereceden polinomsal fonksiyonlar kullanılarak Görünüşte İlişkisiz Regresyon (Seemingly Unrelated Regression-SUR) yöntemi ile modellenmiştir. Girdi değişkenleri ve yanıt değişkenleri arasındaki fonksiyonel ilişkiyi açıklayacak tahmini yanıt modelleri elde

edildikten sonra bu tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleriyle tahmini yanıt modellerinin eş anlı optimizasyonu ile ilgilenilmiştir.

Y_1 ve Y_4 yanıtları için SUR analiziyle elde edilen tahmini yanıt modelleri sırasıyla

$$\hat{Y}_1 = 1.526 - 0.575x_1 - 0.524x_2 - 0.171x_1^2 + 0.318x_1x_2 \quad (6.14)$$

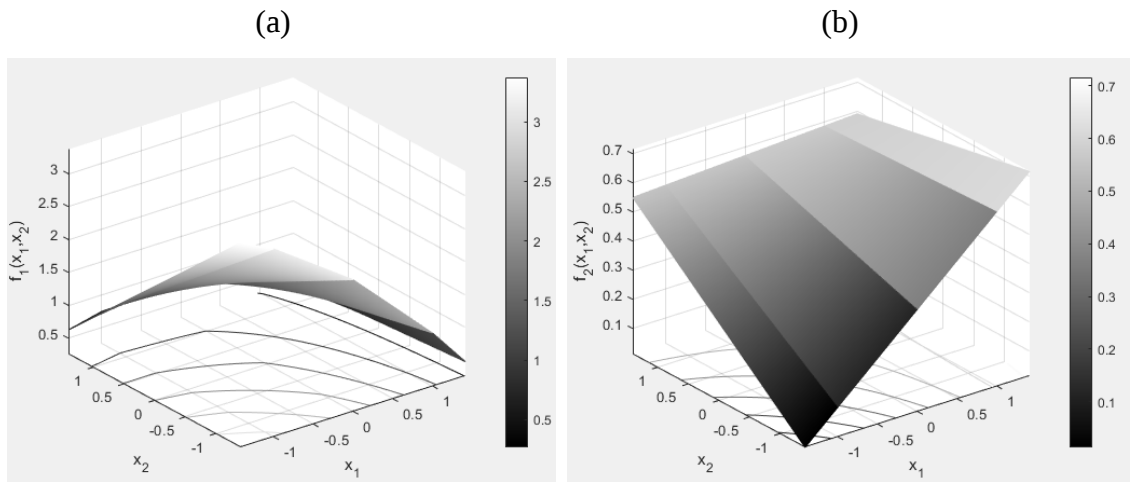
ve

$$\hat{Y}_4 = 0.468 + 0.131x_1 + 0.073x_2 - 0.083x_1x_2 \quad (6.15)$$

dır. Optimal değeri elde edilmek istenilen her bir yanıt fonksiyonu, bir amaç fonksiyonu olarak değerlendirildiğinde

$$\begin{aligned} \max f_1 &= \hat{Y}_1(x) \\ \min f_2 &= \hat{Y}_4(x) \\ x &\in S \end{aligned} \quad (6.16)$$

biçiminde MOO problemine dönüşür. Burada, $\{x_1, x_2\}$ değişkenlerinin aldığı değerler vektörü x olarak belirtilmiştir. f_1 ve f_2 amaç fonksiyonları, sırasıyla $\hat{Y}_1$ ve $\hat{Y}_4$ tahmini yanıt fonksiyonlarına karşılık gelir. Elde edilen tahmini yanıt fonksiyonları için yüzey grafikleri şekil 6.11 (a)-(b)'de görülmektedir. Buna göre, şekil 6.11 (a)'da bir maksimizasyon problemi ve şekil 6.11 (b)'de bir minimizasyon problemi ile ilgilenildiği açıktır.



Şekil 6.11 (a) f_1 ve (b) f_2 amaç fonksiyonları için yüzey grafikleri

Çizelge 6.12-6.13'te, bu çalışmada uygulanan MDNSGA-II ve MDNSGA-III yöntemlerinin ayarlanabilir parametreleri ve parametrelerin aldığı değerler yer almaktadır.

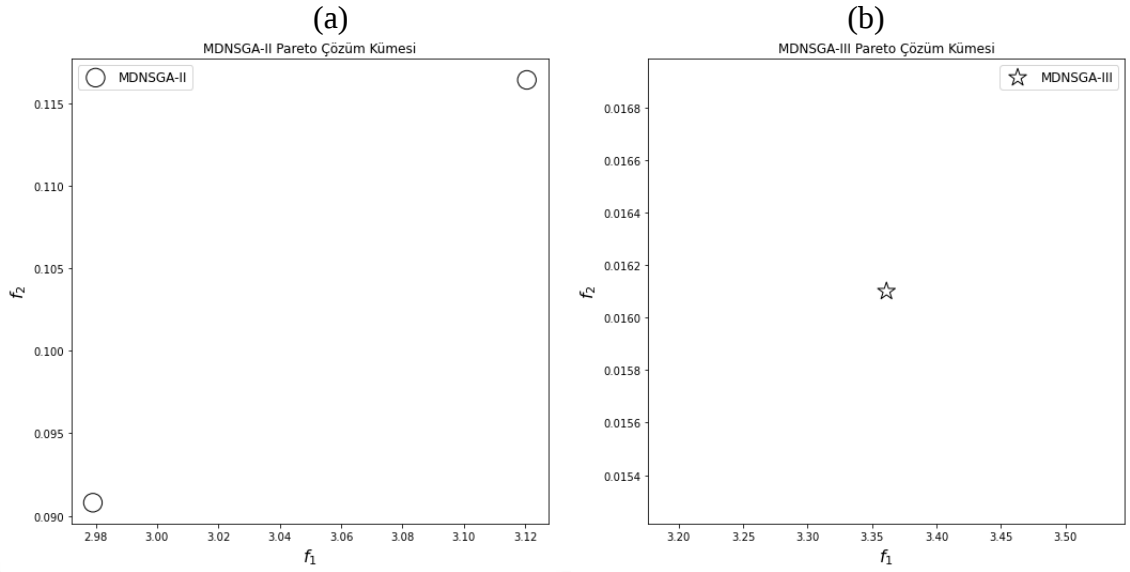
Çizelge 6.12 MDNSGA-II için ayarlanabilir parametreler ve değerleri

<i>Ayarlanabilir parametreler</i>	<i>Parametre değerleri</i>
Girdi değişken sayısı (p)	2
Popülasyon büyüklüğü (N)	50
Yineleme sayısı (n_{gen})	100
Çaprazlama olasılığı (P_c)	0.90
Çaprazlama indeksi (η_c)	20
Mutasyon olasılığı (P_m)	1/2
Mutasyon indeksi (η_m)	20
Çaprazlama operatörü	SBX
Mutasyon operatörü	Polinomsal
Seçim operatörü	Turnuva

Çizelge 6.13 MDNSGA-III için ayarlanabilir parametreler ve değerleri

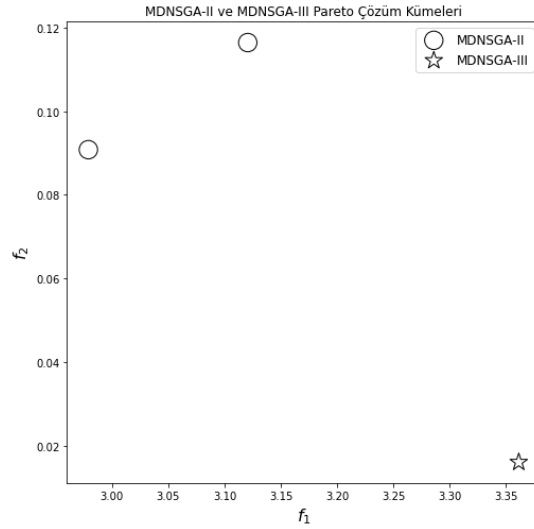
<i>Ayarlanabilir parametreler</i>	<i>Parametre değerleri</i>
Girdi değişken sayısı (p)	2
Popülasyon büyüklüğü (N)	50
Yineleme sayısı (n_{gen})	100
Referans bölüm sayısı (n_{div})	12
Çaprazlama olasılığı (P_c)	0.90
Çaprazlama indeksi (η_c)	20
Mutasyon olasılığı (P_m)	1/2
Mutasyon indeksi (η_m)	20
Çaprazlama operatörü	SBX
Mutasyon operatörü	Polinomsal

Şekil 6.12'de MDNSGA-II ve MDNSGA-III yöntemleriyle elde edilen Pareto çözüm kümeleri yer almaktadır. Buna göre, her iki yöntemle de $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümesi sınırlı sayıda noktaya yakınsamış ve bu nedenle de çözüm çeşitliliği elde edilememiştir.



Şekil 6.12 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümeleri

Şekil 6.13'te, MDNSGA-II ve MDNSGA-III yöntemleriyle elde edilen Pareto çözüm kümeleri birlikte verilmiştir. $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için MDNSGA-III yöntemiyle elde edilen Pareto çözümü daha iyidir.



Şekil 6.13 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması

Çizelge 6.14'te, MDNSGA-II ve MDNSGA-III yöntemleriyle amaç fonksiyonlarını eş anlı optimize eden optimal girdi değişken değerleri yer almaktadır.

Çizelge 6.14 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_1$ ve $\hat{Y}_4$ yanıtları için elde edilen optimal girdi değişken değerleri ve amaç fonksiyonu değerleri

<i>Yöntem</i>	$\mathbf{x} = [x_1 \ x_2]$	$\mathbf{f} = [f_1 \ f_2]$
MDNSGA-II	[-1 -1.414]	[3.1206 0.1164]
	[-1.414 -1.0086]	[2.9791 0.0908]
MDNSGA-III	[-1.414 -1.3641]	[3.3254 0.0231]

y_2 ve y_3 yanıtları için SUR analiziyle elde edilen tahmini yanıt modelleri sırasıyla

$$\hat{Y}_2 = 0.660 - 0.092x_1 - 0.096x_1^2 - 0.058x_2^2 - 0.070x_1x_2 \quad (6.17)$$

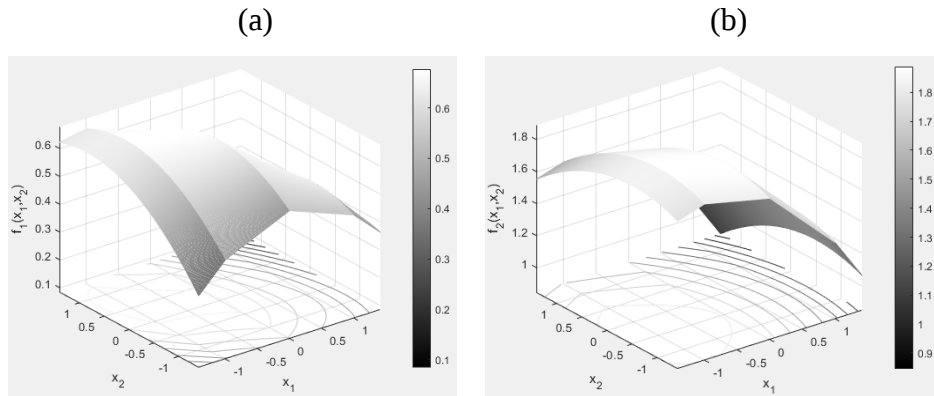
ve

$$\hat{Y}_3 = 1.776 - 0.25x_1 - 0.078x_2 - 0.156x_1^2 - 0.079x_2^2 \quad (6.18)$$

dur. Optimal değeri elde edilmek istenilen her bir yanıt fonksiyonu, bir amaç fonksiyonu olarak değerlendirildiğinde

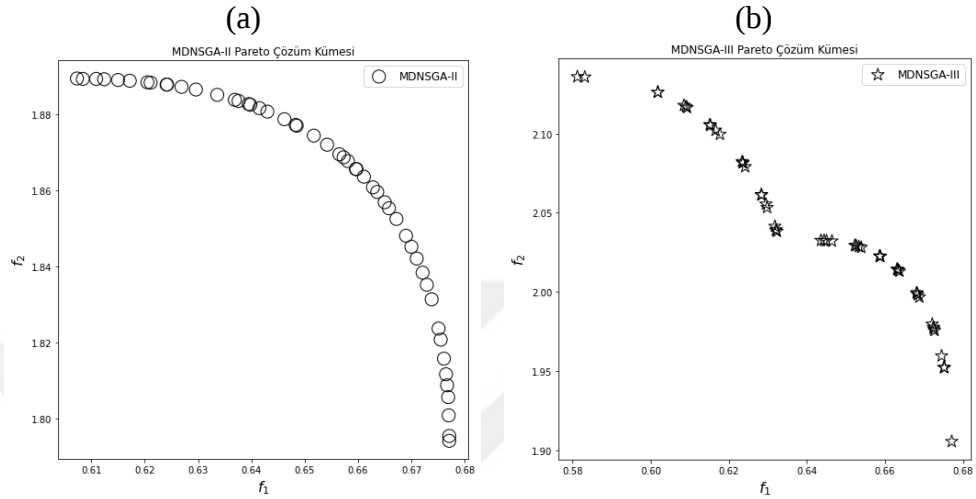
$$\begin{aligned} \max f_1 &= \hat{Y}_2(\mathbf{x}) \\ \max f_2 &= \hat{Y}_3(\mathbf{x}) \\ \mathbf{x} &\in S \end{aligned} \quad (6.19)$$

biçiminde MOO problemine dönüşür. Burada, $\{x_1, x_2\}$ değişkenlerinin aldığı değerler vektörü $\mathbf{x}$ olarak belirtilmiştir. f_1 ve f_2 amaç fonksiyonları, sırasıyla $\hat{Y}_2$ ve $\hat{Y}_3$ tahmini yanıt fonksiyonlarına karşılık gelir. Elde edilen tahmini yanıt fonksiyonları için yüzey grafikleri şekil 6.14 (a)-(b)'de görülmektedir. Şekil 6.14 (a)-(b)'ye bakıldığında maksimizasyon problemleriyle ilgilenildiği görülmektedir.



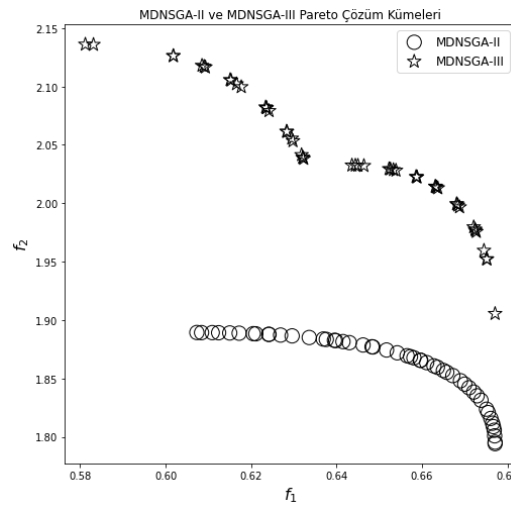
Şekil 6.14 (a) f_1 ve (b) f_2 amaç fonksiyonları için yüzey grafikleri

Çizelge 6.12-6.13'te yer alan MDNSGA-II ve MDNSGA-III yöntemlerinin ayarlanabilir parametre değerleri kullanılarak MOO probleminin MDNSGA-II ve MDNSGA-III ile elde edilen Pareto çözüm kümesi Şekil 6.15'te verilmiştir. Gıda karma veri kümesi için MDNSGA-II ve MDNSGA-III yöntemleriyle elde edilen Pareto çözüm kümeleri EK 4 ve EK 5'te verilmiştir.



Şekil 6.15 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için elde edilen Pareto çözüm kümeleri

Şekil 6.16'da MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması verilmiştir. Şekil 6.16'dan $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için MDNSGA-III'ün daha iyi sonuçlar verdiği görülmektedir.



Şekil 6.16 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_2$ ve $\hat{Y}_3$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması

Y_3 ve Y_4 yanıtları için elde edilen tahmini yanıt modelleri sırasıyla

$$\hat{Y}_3 = 1.776 - 0.25x_1 - 0.078x_2 - 0.156x_1^2 - 0.079x_2^2 \quad (6.20)$$

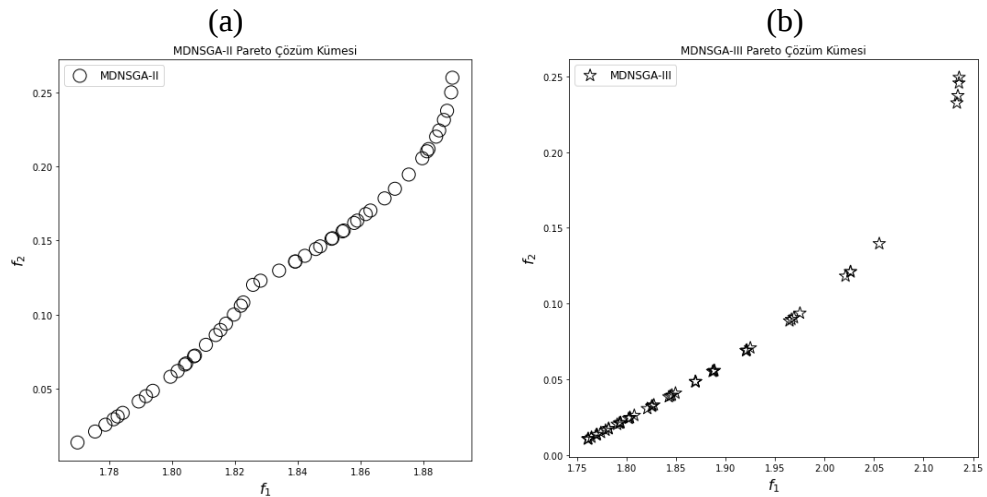
ve

$$\hat{Y}_4 = 0.468 + 0.131x_1 + 0.073x_2 - 0.083x_1x_2 \quad (6.21)$$

dır. Optimal değeri elde edilmek istenilen her bir yanıt fonksiyonu, bir amaç fonksiyonu olarak değerlendirildiğinde

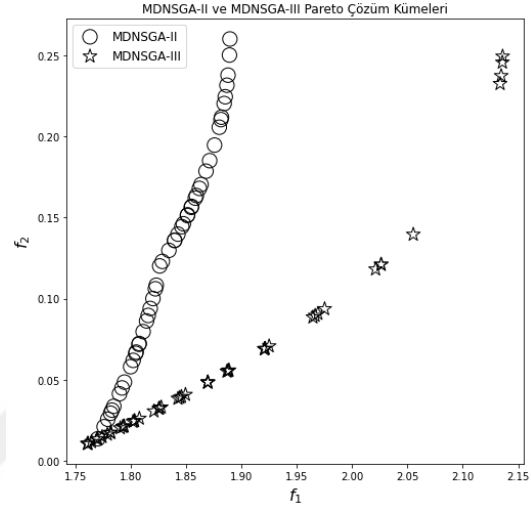
$$\begin{aligned} \max f_1 &= \hat{Y}_3(\mathbf{x}) \\ \min f_2 &= \hat{Y}_4(\mathbf{x}) \\ \mathbf{x} &\in S \end{aligned} \quad (6.23)$$

biçiminde MOO problemine dönüşür. Burada, $\{x_1, x_2\}$ değişkenlerinin aldığı değerler vektörü $\mathbf{x}$ olarak belirtilmiştir. f_1 ve f_2 amaç fonksiyonları, sırasıyla $\hat{Y}_3$ ve $\hat{Y}_4$ tahmini yanıt fonksiyonlarına karşılık gelir. Çizelge 6.12-6.13'te yer alan MDNSGA-II ve MDNSGA-III yöntemlerinin ayarlanabilir parametre değerleri kullanılarak MOO probleminin MDNSGA-II ve MDNSGA-III ile elde edilen Pareto çözüm kümeleri şekil 6.17'de verilmiştir. MDNSGA-II ve MDNSGA-III ile gıda karma veri kümesi için elde edilen Pareto çözüm kümeleri EK 6-7'de verilmiştir.



Şekil 6.17 (a) MDNSGA-II ve (b) MDNSGA-III yöntemleriyle $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümeleri

Şekil 6.18’de MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması verilmiştir. Şekil 6.18’den $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için MDNSGA-III’ün daha iyi sonuçlar verdiği görülmektedir.



Şekil 6.18 MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_3$ ve $\hat{Y}_4$ yanıtları için elde edilen Pareto çözüm kümelerinin karşılaştırılması

7. SONUÇ VE TARTIŞMA

Bu tez çalışmasında, karma veri içeren problemlerde yanıt değişkenlerinin modellenmesi ve elde edilen tahmini yanıt fonksiyonlarının çok amaçlı optimizasyonu ile ilgilenilmiştir. Modelleme aşamasında, GLM, Ridge regresyon, LASSO regresyon ve Elastik Net regresyonu ile SUR analizi uygulanmıştır. Karma veri içeren çok yanıtlı modellerin tahmini yanıt fonksiyonları, amaç fonksiyonları olarak ele alınıp problem MOO problemi biçiminde değerlendirilmiştir. MOO için popülasyon tabanlı bir yapay zeka optimizasyon algoritması olan GA'ya dayalı NSGA-II ve NSGA-III yöntemleri ele alınmıştır. Karma verili çok amaçlı problemlerin optimizasyonu için NSGA-II ve NSGA-III yöntemlerinin karar değişkenleri gösterimi, başlangıç popülasyonunun oluşturulması ve genetik operatörlerin uygulanması aşamalarında uyarlamalar yapılmıştır. Çalışmada, rasyonel değerli kesikli değişkenlerin sıra numaraları dikkate alınarak NSGA-II ve NSGA-III yöntemleri uyarlanmıştır. Uyarlanan yöntemler, bu çalışma kapsamında MDNSGA-II ve MDNSGA-III olarak adlandırılmıştır. Karma veri içeren çok yanıtlı problemler için MDNSGA-II ve MDNSGA-III ile Pareto çözüm kümelerinin elde edilebilir olduğu gösterilmiştir.

UCI veri tabanından enerji verimliliği karma veri kümesi ile literatürde tanınmış gıda alanından deneysel veri kümesi karma veri kümesi haline getirilerek uygulamada kullanılmıştır. Enerji karma veri kümesinde, girdi değişkenleri arasındaki çoklu bağlantı sorunu nedeniyle modelleme aşamasında Ridge, LASSO ve Elastik Net regresyon analizleri yapılmıştır. Elde edilen $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarının tez çalışmasında önerilen MDNSGA-II ve MDNSGA-III yöntemleriyle çok amaçlı optimizasyonu gerçekleştirilmiştir. Bulunan Pareto çözüm kümelerinde MDNSGA-III yöntemiyle çözüm çeşitliliğinin sağlandığı ve amaç fonksiyonlarında iyileşme olduğu gözlenmiştir. Pareto çözüm kümesi, amaç fonksiyonunun yapısına bağlı olduğundan amaç fonksiyonları farklılaştıkça Pareto çözüm kümesindeki değerler de değişir. Enerji verimliliği karma veri kümesindeki yanıtlar için tahmini yanıt fonksiyonları IRLS yöntemiyle elde edilen modeller olarak dikkate alındığında Pareto çözüm kümesi farklılık göstermiştir. Bununla birlikte MDNSGA-III ile MDNSGA-II'ye göre Pareto çözüm kümesinde çözüm çeşitliliğinin sağlanabildiği görülmüştür.

Gıda karma veri kümesinde ise, yanıt değişkenleri arasındaki ilişki dikkate alınarak SUR analiziyle elde edilen modellerin MDNSGA-II ve MDNSGA-III yöntemleriyle çok amaçlı optimizasyonu gerçekleştirilmiştir. MDNSGA-II ve MDNSGA-III yöntemleriyle $\hat{Y}_2 - \hat{Y}_3$ ve $\hat{Y}_3 - \hat{Y}_4$ ikili yanıt fonksiyonları için Pareto çözüm kümesinde çözüm çeşitliliği elde edilirken $\hat{Y}_1 - \hat{Y}_4$ tahmini yanıt fonksiyonları için sınırlı sayıda noktaya yakınsama gözlenmiştir. MDNSGA-III yöntemiyle elde edilen Pareto çözümler daha iyidir. Buna göre, karma verili problemlerin çok amaçlı optimizasyonu için bu çalışmada önerilen MDNSGA-II ve MDNSGA-III yöntemleriyle Pareto çözümlerin elde edilebilir olduğu görülmüştür.



KAYNAKLAR

- Ahmadi, M., Arabi, M., Hoag, D. L. and Engel, B. A. 2013. A Mixed Discrete-Continuous Variable Multiobjective Genetic Algorithm for Targeted Implementation of Nonpoint Source Pollution Control Practices. *Water Resources Research*, 49, 8344–8356.
- Anderson, C. J., Verkuilen, J. and Johnson, T. R. 2012. *Applied Generalized Linear Mixed Models: Continuous and Discrete Data*. Springer, 94 p., New York.
- Apaydın, A. 2002. *Optimizasyon*. AÜFF Döner Sermaye İşletmesi Yayınları, 389 s., Ankara.
- Arora, J. S. and Huang, M. W. 1994. Methods for Optimization Variables: A Review. *Structural Optimization*, 8, 69-85.
- Arora, J. S. 2000. Methods for Discrete Variable Structural Optimization. *Advanced Technology in Structural Engineering*, 1-8.
- Ayan, B., Lorcü F., Erduran, G. Y., Küçükoğlu, M. T., Güçlü, P. ve Abacıoğlu, S. 2021. *Yöneylem Araştırması 2*. Detay Yayıncılık, 346 s., Ankara.
- Babu, B. V. and Angira, R. 2002. A Differential Evolution Approach for Global Optimization of MINLP Problems. In *Proceedings of Fourth Asia Pacific Conference on Simulated Evolution and Learning (SEAL2002)*, 2, Singapore, 880–884.
- Bakır, M. A. ve Altunkaynak B. 2003. *Tamsayılı Programlama Teori, Modeller ve Algoritmalar*. Nobel Yayınevi, 625 s., Ankara.
- Bhesdadiya, R. H., Trivedi, I. N., Jangir, P., Jangir, N. and Kumar, A. 2016. An NSGA-III Algorithm for Solving Multi-Objective Economic/Environmental Dispatch Problem. *Cogent Engineering*, 3(1), 1-23.
- Bolat, B., Erol, K. O. and İmrak, C. 2004. Genetic Algorithms in Engineering Applications and The Function of Operators. *Journal of Engineering and Natural Sciences*, Sigma, 264-271.
- Cebeci, Z. 2021. *R ile Genetik Algoritmalar ve Optimizasyon Uygulamaları*. Nobel Yayınevi, 535 s., Ankara.
- Chaudhari, P., Thakur, A. K., Kumar, R., Banerjee, N. and Kumar, A. 2022. Comparison of NSGA-III with NSGA-II for Multi Objective Optimization of Adiabatic Styrene Reactor. *Materials Today:Proceedings*, 57(4), 1509-1514.

- Chavent, M., Kuentz-Simonet V., Labenne, A. and Saracco, J. 2017. Multivariate Analysis of Mixed Data: The R Package PCAmixdata. arXiv:1411.4911v4 [stat.CO].
- Choi, Y. H. and Kim, J. H. 2019. Self-Adaptive Models for Water Distribution System Design Using Single/Multi-Objective Optimization Approaches. *Water*, 11(6), 1293-1319.
- Ciro, G. C., Dugardin, F., Yalaoui, F. and Kelly, R. 2016. A NSGA-II and NSGA-III Comparison for Solving an Open Shop Scheduling Problem with Resource Constraints. *IFAC-PapersOnLine*, 49(12), 1272-1277.
- Collins, D. 2008. The Performance of Estimation Methods for Generalized Linear Mixed Models. Doctor of Philosophy thesis, University of Wollongong, School of Mathematics and Applied Statistics- Faculty of Informatics, 223, Australia.
- Cooper, M. W. 1981. Survey of Methods for Nonlinear Integer Programming. *Management Science*, 27, 353-361.
- Danish, M., Kumar, S., Qamareen, A. and Kumar, S. 2006. Optimal Solution of MINLP Problems Using Modified Genetic Algorithm. *Chemical Product and Process Modeling*, 1(1), 1-41.
- Das, I. and Dennis, J. E. 1998. Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems. *SIAM Journal on Optimization*, 8(3), 631-657.
- Das, A. K. and Pratihar, D. K. 2018. A Direction-Based Exponential Mutation Operator for Real-Coded Genetic Algorithm. 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT), 12-13 January 2018, 1-4, Kolkata, India.
- De Jong, K. A. 1975. An Analysis of the Behavior of a Class of Genetic Adaptive Systems, Ph.D. Dissertation, Department of Computer and Communication Science, University of Michigan, Ann Arbor, MI.
- Deb, K. 2001. Multi-Objective Optimization Using Evolutionary Algorithms, 1st ed, John Wiley-Interscience series in systems and optimization, 498 p.
- Deb, K. 2012. Optimization for Engineering Design Algorithms and Examples, Second Edition, Phi Learning Private Limited, New Delhi.
- Deb, K. and Agrawal, S. 1999. A Niched-Penalty Approach for Constraint Handling in Genetic Algorithms. *Artificial Neural Nets and Genetic Algorithms*, 235–243.
- Deb, K. and Jain, H. 2014. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving

- Problems with Box Constraints. IEEE Transactions on Evolutionary Computation, 18(4), 577-601.
- Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2),182-197.
- Deep, K., Singh, K. P., Kansal, M. L. and Mohan, C. 2009. A Real Coded Genetic Algorithm for Solving Integer and Mixed Integer Optimization Problems. Applied Mathematics and Computation, 212, 505-518.
- Deep, K. and Thakur, M. 2007a. A New Crossover Operator for Real Coded Genetic Algorithms. Applied Mathematics and Computation, 188, 895-911.
- Deep, K. and Thakur, M. 2007b. A New Mutation Operator for Real Coded Genetic Algorithms. Applied Mathematics and Computation, 193, 211-230.
- Ebrahimi, A., Tamnanloo, J., Mousavi, S. H., Soroodan, M. E., Hosseini, E., Ghasemi, H. and Mozaffari, S. 2021. Discrete-Continuous Genetic Algorithm for Designing a Mixed Refrigerant Cryogenic Process. Industrial & Engineering Chemistry Research, 60(20), 7700-7711.
- Eğrioğlu, E., Yolcu, U. ve Baş, E. 2019. Yapay Sinir Ağları Öngörü ve Tahmin Uygulamaları. Nobel Yayınevi, 252 s., Ankara.
- El-Kribi, B., Houidi, A., Affi, Z. and Romdhane, L. 2013. Application of Multi-Objective Genetic Algorithms to The Mechatronic Design of a Four Bar System with Continuous and Discrete Variables. Mechanism and Machine Theory, 61, 68–83.
- Emel, G. G. ve Taşkın, Ç. 2002. Genetik Algoritmalar ve Uygulama Alanları. Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, XXI, 1, 129-152.
- Ergül, E. U. 2010. Çok Amaçlı Genetik Algoritmalar: Temelleri ve Uygulamaları. Doktora Tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 217, Samsun.
- Faraway, J. J. 2006. Extending the Linear Model with R Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman & Hall/CRC Taylor & Francis Group.
- Floudas, C. A. 1995. Nonlinear Mixed-Integer Optimization. Fundamentals and Applications Oxford University Press, New York, USA.
- Fonseca, C. M. and Fleming, P. J. 1993. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. In Proceedings of the Fifth International Conference on Genetic Algorithms, S. Forrest, Ed. San Mateo, CA: Morgan Kauffman, 416–423.

- Fox, J. 2015. Applied Regression Analysis and Generalized Linear Models. SAGE Publications, Third Edition, 816 p., Canada.
- Friendly, M. and Meyer, D. 2015. Discrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data. Chapman and Hall/CRC Published.
- Gambier, A. 2008. MPC and PID Control Based on Multi-Objective Optimization. 2008 American Control Conference Westin Seattle Hotel, Seattle, Washington, USA June 11-13.
- Garrido, J. and Zhou, J. 2009. Full Credibility with Generalized Linear and Mixed Models. ASTIN Bulletin, 39(1), 61-80.
- Grefenstette, J. J. 1986. Optimization of Control Parameters for Genetic Algorithms. IEEE Transactions on Systems, Man and Cybernetics, 16(1), 122-128.
- Grossmann, I. E. 2002. Review of Non-Linear Mixed Integer and Disjunctive Programming Techniques. Optimization and Engineering, 3, 227-252.
- Ghasemi, M. R., Hinton, E. and Wood, R. D. 1997. Optimization of Trusses Using Genetic Algorithms for Discrete and Continuous Variables. Engineering Computations, 16(3), 272-301.
- Gu, Q., Wang, R., Xie, H., Li, X., Jiang, S. and Xiong, N. 2021. Modified Non-Dominated Sorting Genetic Algorithm III with Fine Final Level Selection. Applied Intelligence, 51, 4236–4269.
- Guangyong, S., Huile, Z., Jianguang, F., Guangyao, L. and Qing, L. 2018. A New Multi-Objective Discrete Robust Optimization Algorithm for Engineering Design. Applied Mathematical Modelling, 53, 602-621.
- Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, ABD.
- Hajela, P. and Shih, C. J. 1989. Multiobjective Optimum Design in Mixed Integer and Discrete Design Variable Problems. AIAA Journal, 28(4), 670-678.
- Hoerl, A. E. and Kennard, R. W. 1970a. Ridge Regression: Applications to Non-Orthogonal Problems. Technometrics, 12(1), 69-82.
- Hoerl, A. E. and Kennard, R. W. 1970b. Ridge Regression: Biased Estimation for Non-Orthogonal Problems. Technometrics, 12(1), 55-67.
- Holland, J. 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, ABD.

- Holzmann, T. and Smith, J. C. 2018. Solving Discrete Multi-Objective Optimization Problems Using Modified Augmented Weighted Tchebychev Scalarizations. *European Journal of Operational Research*, 271, 436–449.
- Horn, J., Nafploitis, N. and Goldberg, D. E. 1994. A Niche Pareto Genetic Algorithm for Multiobjective Optimization. In *Proceedings of the First IEEE Conference on Evolutionary Computation*, Z. Michalewicz, Ed. Piscataway, NJ: IEEE Press, 82–87.
- Huang, M. W. and Arora, J. S. 1995. *Engineering Optimization with Discrete Variables*. American Institute of Aeronautics and Astronautics, 95, 1475-1485.
- Ishibuchi, H., Imada, R., Setoguchi, Y. and Nojima, Y. 2016. Performance Comparison of NSGA-II and NSGA-III on Various Many-Objective Test Problems. 2016 IEEE Congress on Evolutionary Computation (CEC), 3045-3052, Vancouver, BC, Canada.
- Ishibuchi, H. and Shibata, Y. 2004. Mating Scheme for Controlling the Diversity-Convergence Balance for Multiobjective Optimization, *Proceedings of 2004 Genetic and Evolutionary Computation Conference, Part I*, Seattle, ABD, 1259-1271.
- İşçi, Ö. ve Korukoğlu, S. 2003. Genetik Algoritma Yaklaşımı ve Yöneylem Araştırmasında Bir Uygulama. *Yönetim ve Ekonomi Dergisi*, 10(2), 191-208.
- Jaber, A., Lafon, P. and Younes, R. 2022. A Branch and Bound Algorithm Based on NSGA-II for Multi-Objective Mixed Integer Nonlinear Optimization Problems. *Engineering Optimization*, 54(6), 1004–1022.
- James, G., Witten, D., Hastie, T., Tibshirani, R. and Taylor, T. 2023. *An Introduction to Statistical Learning with Applications in Python*, Chapter 6: Linear Model Selection and Regularization. Allen, G., Veaux, R. D. and Nugent, R. (eds), Springer Texts in Statistics, 229-288, USA.
- Kanata, S., Suwarno, S., Sianipar, G. H. and Maulidevi, N. U. 2019. Non-dominated Sorting Genetic Algorithm-III for Multi-objective Optimal Reactive Power Dispatch Problem in Electrical Power System. 2019 2nd International Conference on High Voltage Engineering and Power Systems (ICHVEPS), Bali, Indonesia.
- Karaboğa, D. 2020. *Yapay Zeka Optimizasyon Algoritmaları*. Nobel Yayınevi, 246 s., Ankara.
- Karakoç, G. ve Türkşen, Ö. 2023. Karma Veri İçeren Çok Yanıtlı Problemlerde NSGA-II'ye Dayalı Bir Optimizasyon Yaklaşımı. *İstatistikçiler Dergisi: İstatistik & Aktüerya*, 16(2), 57-80.
- Katoch, S., Chauhan, S. S. and Kumar, V. 2021. A Review on Genetic Algorithm: Past, Present and Future. *Multimedia Tools and Applications*, 80, 8091–8126.

- Kaya, S. ve Fırlalı, N. 2016. Çok Amaçlı Optimizasyon Problemlerinde Pareto Optimal Kullanımı. Social Sciences Research Journal, 5(2), 9-18, ISSN:2147-5237.
- Keskintürk, T. 2006. Diferansiyel Gelişim Algoritması. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 5(9), 85-99.
- Khuri, A. I. and Cornell, M. 1996. Response Surfaces, Marcel Dekker, Inc. 510 p., NewYork.
- Khuri, A. I., Mukherjee, B., Sinha, B. K. and Ghosh, M. 2006. Design Issues for Generalized Linear Models: A Review. Statistical Science, 21(3), 376-399.
- Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. 1983. Optimization by Simulated Annealing. Science, 220, 671-680.
- Kondoh, T., Tatsukawa, T., Oyama, A. Takeshi, W. and Fujii, K. 2016. Effects of Discrete Design-Variable Precision on Real-Coded Genetic Algorithm. 2016 IEEE Symposium Series on Computational Intelligence (SSCI).
- Kurtuluş, M. 2001. Ridge Regresyon Üzerine Bir Çalışma. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri, İstatistik Anabilim Dalı, Ankara.
- Küçükkoç, İ. 2020. EMM3208 Optimizasyon Teknikleri Ders Notları. Balıkesir Üniversitesi, Endüstri Mühendisliği.
- Li, Y. and Gen M.,1996. Nonlinear Mixed Integer Programming Problems Using Genetic Algorithm and Penalty Function. 1996 IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems, 4, 2677-2682.
- Lin, C. Y. and Hajela, P. 1992. Genetic Algorithms in Optimization Problems with Discrete and Integer Design Variables. Engineering Optimization, 19(4), 309- 327.
- Lin, S. S., Zhang, C. C. and Wang, H. P. B. 1995. On Mixed-Discrete Nonlinear Optimization Problems: A Comparative Study. Engineering Optimization, 23(4), 287-300.
- Lorcu, F. 2016. Yöneylem Araştırması. Detay Yayıncılık, 260 s., Ankara.
- Marchand, H., Martin, A. and Weismantel, R. 2002. Cutting Planes in Integer and Mixed Integer Programming. Discrete Applied Mathematics, 123, 397-446.
- McCullagh, P. and Nelder, J. A., 1989. Generalized Linear Models. Second Edition. Chapman and Hall/CRC, 511 p., London.
- Merah, A. and Adjabi, M. 2018. Multi-Objective Optimization in Power Systems Including UPFC Controller with NSGA-III. International Journal of Modeling and Optimization, 8(6), 318-325.

- Michalewicz, Z. 1994. Genetic Algorithms + Data Structures = Evolutionary Programs. 2nd Edition, Springer-Verlag, 340 p., Berlin.
- Miettinen, K. 1998. Nonlinear Multiobjective Optimization. Kluwer Academic Publishers, 10 p., USA.
- Montgomery, D., Peck, E. A. and Vining, G. G. 2012. Introduction to Linear Regression Analysis. Beşinci Basımdan Çeviri. Erar, M. A. 2013. Doğrusal Regresyon Analizine Giriş. Nobel Yayınevi, 646 s., Ankara.
- Nelder, J. A. and Wedderburn, R. W. M. 1972. Generalized Linear Models. Journal of the Royal Statistical Society A-General, 135, 370–384.
- Ogut, J. O., Streeck, T. S. and Piepho, H. P. 2012. Genomic Selection Using Regularized Linear Regression Models: Ridge Regression, Lasso, Elastic Net and Their Extensions. BMC Proceedings, 6(2), 1-6.
- Ostertagová, E. 2012. Modelling Using Polynomial Regression. Procedia Engineering, 48, 500 – 506.
- Örkcü, H. H., Özsoy, V. S. ve Koçak, E. 2024. Yöneylem Araştırması. Seçkin Yayıncılık, 202 s., Ankara.
- Öztürk, F. 2012. Genelleştirilmiş Lineer Modeller. Fikri Öztürk’ün Web Sayfası: <http://80.251.40.59/science.ankara.edu.tr/ozturk/Dersler/ist402/Ders6/Ders6.pdf>. Erişim Tarihi:15.11.2021.
- Pichler, M. and Hartig, F. 2022. Machine Learning and Deep Learning – A Review for Ecologists. Methods in Ecology and Evolution, 14(4), 994-1016.
- Rainer, M. 2012. A Genetic Algorithm for Mixed-Integer Multicriteria Optimization Problems and its Application to Engines in order to Optimize Fuel Consumption and Driving Performance. Graz University of Technology, Institute of Mathematics, Master Programme Mathematical Computer Science, Master Thesis, Austria.
- Rajeev, S. and Krishnamoorthy, C. S. 1992. Discrete Optimization of Structures Using Genetic Algorithms. Journal of Structural Engineering, 118(5), 1233-1250.
- Rao, S. S. 2020. Engineering Optimization Theory and Practice. 5th Edition, John Wiley & Sons, USA.
- Rao, S. S. and Xiong, Y. 2005a. A Hybrid Genetic Algorithm for Mixed-Discrete Design Optimization. ASME Journal of Mechanical Design, 127, 1100–1112.
- Rao, S. S. and Xiong, Y. 2005b. Mixed-Discrete Fuzzy Multiobjective Programming for Engineering Optimization Using Hybrid Genetic Algorithm. AIAA Journal, 43(7), 1580-1590.

- Ranstam, J. and Cook, J. A. 2018. LASSO Regression. *BJS Society Statistical nugget*, 105, 1338–1348.
- Roy, S., Crossley, W. A. and Jain, S. 2021. A Hybrid Approach for Solving Constrained Multi-Objective Mixed-Discrete Nonlinear Programming Engineering Problems, In: *Engineering Problems-Uncertainties, Constraints and Optimization Techniques*. Tsuzuki, M. S. G., Takimoto, R. Y., Sato, A., Saka, T., Barari, A. and Rahman, R. O. A. (eds), IntechOpen.
- Santos, F., Costa, L. and Varela, L. 2024. Performance Comparison of NSGA-II and NSGA-III on Bi-objective Job Shop Scheduling Problems, In: *Pereira, A.I., Mendes, A., Fernandes, F. P., Pacheco, M. F., Coelho, J. P., Lima, J. (eds) Optimization, Learning Algorithms and Applications*, 531-543, Springer, Cham.
- Satman, M. H. 2016. *Genetik Algoritmalar*. Türkmen Kitabevi, İstanbul.
- Salkin, H. M. 1975. *Integer Programming*. Eddison Wesley Publishing Com, Amsterdam.
- Schaffer, J. D., 1984. Some Experiments in Machine Learning Using Vector Evaluated Genetic Algorithms. Ph.D. Thesis, Vanderbilt University, Nashville, ABD.
- Schaffer, J. D., Caruana, R. A., Eshelman, L. J. and Das, R. 1989. A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization. *Proceedings of the 3rd International Conference on Genetic Algorithms*, George Mason University, Fairfax, Virginia, USA.
- Schmidt, R. H., Illingworth, B. L., Deng, J. C. and Cornell, J. A. 1979. Multiple Regression and Response Surface Analysis of the Effects of Calcium Chloride and Cysteine on Heat-Induced Whey Protein Gelation. *Journal of Agricultural and Food Chemistry*, 27(3), 529-532.
- Seada, H. and Deb, K. 2015. U-NSGA-III: A Unified Evolutionary Algorithm for Single, Multiple and Many-Objective Optimization. *Conference: International Conference on Evolutionary Multi-Criterion Optimization*, 1-30.
- Sen, A. and Srivastava, M. 1990. *Regression Analysis: Theory, Methods and Application*. 1st Edition, Springer-Verlag, 364 p., New York.
- Song, S., Jin, H. and Yang, Q. 2021. Performance Analysis of Different Operators in Genetic Algorithm for Solving Continuous and Discrete Optimization Problems, In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021)*, 1, 536-547.
- Srinivas, N. and Deb, K. 1994. Muultiobjective Optimization Using Non-Dominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3), 221-248.
- Tanabe, R. and Ishibuchi, H. 2020. An Easy to Use Real Word Multi Objective Optimization Problem Suite. *Applied Soft Computing Journal*, 89, 1-9.

- Taştan, S. 2012. Genetik Algoritma Temelinde Rejim Değişikliği Modellerinin Belirlenmesi. Doktora Tezi, Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Anabilim Dalı, Sayısal Yöntemler Bilim Dalı, Sivas.
- Tavana, M., Li, Z., Mobin, M., Komaki, M. and Teymourian, E. 2016. Multi-Objective Control Chart Design Optimization Using NSGA-III and MOPSO Enhanced with DEA and TOPSIS. Expert Systems with Applications, 50, 17-39.
- Tibshirani, R. 1996. Regression Shrinkage and Selection via the LASSO. Journal of Royal Statistical Society B, 58, 267–288.
- Tsanas, A. and Xifara, A. 2012. Accurate Quantitative Estimation of Energy Performance of Residential Buildings Using Statistical Machine Learning Tools. Energy and Buildings, 49, 560-567.
- Tong, W., Chowdhury, S. and Messac, A. 2014. A new multi-Objective Mixed-Discrete Particle Swarm Optimization Algorithm. Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference Buffalo, New York, USA.
- Tunçel, S. 2022. Çok Yanıtlı Deneysel Verilerin Görünüşte İlişkisiz Regresyon Analizi ile Modellenmesi ve Optimal Değişken Değerlerinin Belirlenmesi. Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Ankara.
- Türkşen, Ö. 2011. Çok Yanıtlı Yüzey Problemlerinin Çözümüne Bulanık ve Sezgisel Yaklaşım. Doktora Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Ankara.
- Türkşen, Ö., Tunçel, S. and Vural, N. 2019. A Seemingly Unrelated Regression Modeling for Extraction Process in Green Chemistry. y-BIS Conference 2019: Recent Advances in Data Science and Business Analytics, September 25–28, Mimar Sinan Fine Arts University, Proceeding Book, 129-133, İstanbul.
- Türkşen, Ö. 2023. Optimizasyon Yöntemleri ve Matlab, Python, R Uygulamaları. Nobel Yayınevi, 448 s., Ankara.
- UCI Machine Learning Repository. Available online: <https://archive.ics.uci.edu/dataset/242/energy+efficiency>. Erişim tarihi: 15.11.2021.
- Uzgören, N. 1999. Görünüşte İlişkisiz Regresyon Denklemlerinin Etkinliği Üzerinde Hatalar Arasındaki Korelasyonun ve Bağımsız Değişkenler Arasındaki Korelasyonun Etkisinin İncelenmesi. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 1, 139-154.
- Vergili, M. 2023. Doğrusal Regresyonda Ridge, Lasso ve Elastik Net Yöntemlerinin Sağlık Alanında Uygulanması. Yüksek Lisans Tezi, Süleyman Demirel

Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Isparta.

- Wang, W., Zmeureanu, R. and Rivard, H. 2005. Applying Multi-Objective Genetic Algorithms in Green Building Design Optimization. *Building and Environment*, 40, 1512–1525.
- Wasanapradit, T., Mukdasanit, N., Chaiyaratana, N. and Srinophakun, T. 2011. Solving Mixed-Integer Nonlinear Programming Problems Using Improved Genetic Algorithms. *Korean Journal of Chemical Engineering*, 28(1), 32-40.
- Wu, S. J. and Chow, P. T. 1994. Genetic Algorithms for Solving Mixed-Discrete Optimization Problems. *Journal of the Franklin Institute*, 331(4), 381-401.
- Wu, S. J. and Chow, P. T. 1995a. Integrated Discrete and Configuration Optimization of Trusses Using Genetic Algorithms. *Computers & Structures*, 55(4), 695-702.
- Wu, S. C. and Chow, P. T. 1995b. Genetic Algorithms for Nonlinear Mixed Discrete-Integer Optimization Problems via Meta-Genetic Parameter Optimization. *Engineering Optimization*, 24(2), 137-159.
- Xiong, Y. and Rao, S. S. 2004. Fuzzy Nonlinear Programming for Mixed-Discrete Design Optimization Through Hybrid Genetic Algorithm. *Fuzzy Sets and Systems*, 146, 167 – 186.
- Yan, L., Shen, K. and Hu, S. 2004. Solving Mixed Integer Nonlinear Programming Problems with Line-up Competition Algorithm. *Computers and Chemical Engineering*, 28, 2647–2657.
- Yang, X. S. 2010. *Engineering Optimization: An Introduction with Metaheuristic Applications*. John Wiley & Sons, 18 p.
- Yannibelli, V., Pacini, E., Monge, D., Mateos, C. and Rodriguez, G. 2020. A Comparative Analysis of NSGA-II and NSGA-III for Autoscaling Parameter Sweep Experiments in the Cloud. *Hindawi Scientific Programming*, 1-17.
- Yiqing, L., Xigang, Y. and Yongjian, L. 2007. An Improved PSO Algorithm for Solving Non-Convex NLP MINLP Problems with Equality Constraints. *Computers and Chemical Engineering*, 31, 153–162.
- Yuan, Y., Xu, H. and Wang, B. 2014. An improved NSGA-III Procedure for Evolutionary Many-Objective Optimization. In: *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, ACM, New York, USA, 661-668.
- Zellner, A. 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias. *Journal of the American Statistical Association*, 57, 348-368.

Zitzler, E., Deb, K. and Thiele, L. 2000. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evolutionary Computation*, 8(2), 173–195.

Zou, H. and Hastie, T. 2005. Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301-320.



EKLER

EK 1 Enerji karma verisi için $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

EK 2 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi

EK 3 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

EK 4 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi

EK 5 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

EK 6 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi

EK 7 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

EK 1 Enerji karma verisi için $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
1	0.973976	563.7122	251.1905	217.1654	3.5	2	0	0	1.3936	4.603
2	0.973951	621.6016	251.1691	211.2734	3.5	2	0	0	1.242	5.5045
3	0.973953	540.2497	251.1711	217.9534	3.5	2	0	0	1.486	4.2909
4	0.973952	745.9243	251.1703	198.3091	3.5	2	0	0	0.9229	7.4503
5	0.973945	604.1705	251.1645	213.0953	3.5	2	0	0	1.2866	5.2316
6	0.974066	514.5	251.2672	220.5	3.5	2	0	0	1.5559	3.8915
7	0.973956	790.0718	251.1739	194.1473	3.5	2	0	0	0.801	8.1267
8	0.973952	796.805	251.1704	193.427	3.5	2	0	0	0.7841	8.2327
9	0.973951	753.6067	251.1691	197.6571	3.5	2	0	0	0.9003	7.5657
10	0.973976	716.86	251.1906	201.2665	3.5	2	0	0	0.9992	6.9977
11	0.97394	769.3332	251.1597	196.2309	3.5	2	0	0	0.8556	7.8048
12	0.973952	584.9241	251.1705	215.0058	3.5	2	0	0	1.3379	4.9335
13	0.973942	802.8802	251.1619	192.814	3.5	2	0	0	0.768	8.3272
14	0.973939	776.6987	251.1596	195.4719	3.5	2	0	0	0.8365	7.9198
15	0.974065	696.1614	251.2663	202.9693	3.5	2	0	0	1.0621	6.6881
16	0.973952	666.5593	251.1703	206.8005	3.5	2	0	0	1.1224	6.201
17	0.973945	706.6423	251.1644	201.9396	3.5	2	0	0	1.0327	6.851
18	0.973944	679.1522	251.1638	205.4786	3.5	2	0	0	1.0902	6.3985
19	0.973952	762.1268	251.1703	196.8173	3.5	2	0	0	0.8775	7.6974
20	0.973944	736.7734	251.1638	199.2065	3.5	2	0	0	0.9474	7.3091
21	0.973943	652.6288	251.1623	208.2565	3.5	2	0	0	1.158	5.9829
22	0.973943	783.3849	251.1625	194.8101	3.5	2	0	0	0.8187	8.0232
23	0.973945	726.5869	251.1642	200.2625	3.5	2	0	0	0.9737	7.1498
24	0.973963	637.9017	251.1793	209.6615	3.5	2	0	0	1.1985	5.7566
25	0.973943	808.5	251.1623	192.2651	3.5	2	0	0	0.7528	8.4139
26	0.973952	761.1592	251.1706	196.9129	3.5	2	0	0	0.8801	7.6824
27	0.973953	745.6015	251.1709	198.3393	3.5	2	0	0	0.9238	7.4454
28	0.973951	603.674	251.1692	213.1472	3.5	2	0	0	1.2879	5,2237
29	0.973946	696.4113	251.1655	202.8974	3.5	2	0	0	1.0611	6.6944
30	0.973942	777.2952	251.1619	195.4105	3.5	2	0	0	0.835	7.9291
31	0.973941	685.5361	251.161	204.898	3.5	2	0	0	1.0721	6.4956
32	0.973945	705.7645	251.1644	202.0425	3.5	2	0	0	1.0348	6.8369
33	0.973945	672.6983	251.1644	206.1304	3.5	2	0	0	1.1072	6.2981
34	0.973946	756.4151	251.1654	197.4315	3.5	2	0	0	0.8917	7.6074
35	0.973943	721.0236	251.1629	200.8235	3.5	2	0	0	0.9883	7.0634
36	0.973946	563.3387	251.1654	217.2042	3.5	2	0	0	1.3943	4.5974
37	0.974066	514.5	251.2675	220.5	3.5	2	0	0	1.5559	3.8915
38	0.973941	651.5444	251.161	208.3964	3.5	2	0	0	1.1602	5.9651
39	0.973952	621.9384	251.17	211.2381	3.5	2	0	0	1.2411	5.5098
40	0.97394	636.9138	251.1602	209.7644	3.5	2	0	0	1.2009	5.7413

EK 1 Enerji karma verisi için $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
41	0.973978	739.6457	251.1916	198.9014	3.5	2	0	0	0.9405	7.3539
42	0.973956	789.8351	251.1738	194.1699	3.5	2	0	0	0.8017	8.123
43	0.973944	786.0465	251.1638	194.5533	3.5	2	0	0	0.8115	8.0642
44	0.973953	765.9987	251.1711	196.4485	3.5	2	0	0	0.8669	7.7568
45	0.973942	581.559	251.1615	215.3441	3.5	2	0	0	1.3467	4.8813
46	0.973938	796.8359	251.1584	193.4228	3.5	2	0	0	0.7839	8.2333
47	0.973939	541.7261	251.1588	217.8412	3.5	2	0	0	1.4812	4.3127
48	0.973961	802.753	251.1769	192.8315	3.5	2	0	0	0.7684	8.3249
49	0.973943	808.0469	251.1626	192.3101	3.5	2	0	0	0.754	8.4069
50	0.973943	730.8521	251.1629	199.8409	3.5	2	0	0	0.9623	7.2158
51	0.974067	807.2604	251.2679	192.3946	3.5	2	0	0	0.7573	8.3935
52	0.973953	758.0157	251.1711	197.2439	3.5	2	0	0	0.8881	7.6331
53	0.973944	678.5907	251.1638	205.5509	3.5	2	0	0	1.0914	6.3892
54	0.973976	717.2719	251.1901	201.2239	3.5	2	0	0	0.9981	7.0041
55	0.973942	793.8531	251.1615	193.748	3.5	2	0	0	0.7913	8.1861
56	0.973959	756.4151	251.1754	197.4315	3.5	2	0	0	0.8918	7.6073
57	0.973946	620.6544	251.1654	211.3686	3.5	2	0	0	1.2444	5.4898
58	0.973943	705.7645	251.1623	202.0425	3.5	2	0	0	1.0348	6.8369
59	0.973977	738.354	251.1913	199.0377	3.5	2	0	0	0.9438	7.3336
60	0.974062	521.5872	251.2637	219.8477	3.5	2	0	0	1.536	3.9996
61	0.973943	786.4158	251.1625	194.5074	3.5	2	0	0	0.8107	8.0702
62	0.973938	541.4111	251.1586	217.8616	3.5	2	0	0	1.4823	4.3082
63	0.973975	727.6937	251.1892	200.1408	3.5	2	0	0	0.9713	7.1671
64	0.973942	581.559	251.1615	215.3441	3.5	2	0	0	1.3467	4.8813
65	0.973942	743.643	251.162	198.5303	3.5	2	0	0	0.929	7.4153
66	0.973952	789.8351	251.1703	194.1699	3.5	2	0	0	0.8016	8.1231
67	0.973938	651.0135	251.1587	208.451	3.5	2	0	0	1.1616	5.9568
68	0.973942	777.4673	251.1618	195.3939	3.5	2	0	0	0.8346	7.9317
69	0.973946	770.7828	251.1651	196.0326	3.5	2	0	0	0.8529	7.829
70	0.973941	661.2098	251.161	207.2343	3.5	2	0	0	1.1384	6.1215
71	0.973949	601.1791	251.1675	213.4	3.5	2	0	0	1.2944	5.185
72	0.973942	633.2862	251.162	210.1327	3.5	2	0	0	1.2104	5.6849
73	0.973942	802.8802	251.1619	192.814	3.5	2	0	0	0.768	8.3272
74	0.973946	698.5562	251.1654	202.6904	3.5	2	0	0	1.0553	6.7275
75	0.973945	573.367	251.1641	215.7889	3.5	2	0	0	1.3757	4.7666
76	0.973944	634.3716	251.163	210.0395	3.5	2	0	0	1.2072	5.7012
77	0.973953	758.0157	251.1711	197.2439	3.5	2	0	0	0.8881	7.6331
78	0.973976	716.86	251.1906	201.2665	3.5	2	0	0	0.9992	6.9977
79	0.973943	653.3822	251.1623	208.1847	3.5	2	0	0	1.1559	5.9945
80	0.97396	726.0496	251.1763	200.3001	3.5	2	0	0	0.9756	7.1419

EK 1 Enerji karma verisi için $\hat{Y}_1^R$ ve $\hat{Y}_2^{EN}$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
81	0.973944	738.7395	251.1635	199.0037	3.5	2	0	0	0.9423	7.3397
82	0.973943	514.5	251.163	220.5	3.5	2	0	0	1.5547	3.8925
83	0.973944	771.585	251.1635	195.9557	3.5	2	0	0	0.8507	7.8414
84	0.973946	708.0802	251.1651	201.783	3.5	2	0	0	1.0292	6.8737
85	0.973963	541.7261	251.1788	217.8412	3.5	2	0	0	1.4814	4.3125
86	0.973938	598.6211	251.1584	213.7005	3.5	2	0	0	1.3002	5.1439
87	0.973978	572.4816	251.1918	215.8788	3.5	2	0	0	1.3784	4.7526
88	0.973946	754.2541	251.165	197.7187	3.5	2	0	0	0.896	7.5715
89	0.97396	687.9774	251.1766	204.5896	3.5	2	0	0	1.0671	6.5354
90	0.973945	669.5264	251.1641	206.4796	3.5	2	0	0	1.1149	6.2479
91	0.973973	792.8767	251.1879	193.8089	3.5	2	0	0	0.7949	8.1719
92	0.973952	584.9241	251.1705	215.0058	3.5	2	0	0	1.3379	4.9335
93	0.973942	743.1953	251.162	198.5735	3.5	2	0	0	0.9302	7.4084
94	0.973942	803.1061	251.1615	192.7956	3.5	2	0	0	0.7673	8.3305
95	0.973939	777.8018	251.159	195.3697	3.5	2	0	0	0.8335	7.9366
96	0.973941	686.0529	251.1609	204.7868	3.5	2	0	0	1.0719	6.5056
97	0.973952	808.0469	251.1701	192.3101	3.5	2	0	0	0.7541	8.4068
98	0.973963	786.3041	251.1793	194.5283	3.5	2	0	0	0.811	8.068
99	0.973946	620.7996	251.165	211.3526	3.5	2	0	0	1.2441	5.4921
100	0.97396	799.3164	251.1766	193.2002	3.5	2	0	0	0.777	8.2707

**EK 2 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin
MDNSGA-II ile elde edilen Pareto çözüm kümesi**

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
1	0.62	553.1395	356.1455	148.7025	3.5	2	0.4	5	10.3660	87.1077
2	0.62	553.8913	355.3345	148.8426	3.5	2	0.4	5	10.3466	87.6671
3	0.62	553.2702	356.1712	148.7327	3.5	2	0.4	5	10.3643	87.1115
4	0.62	553.4816	355.8946	148.7113	3.5	2	0.4	5	10.3583	87.2960
5	0.62	553.4573	355.8442	148.7079	3.5	2	0.4	5	10.3582	87.3209
6	0.62	553.3854	355.9131	148.7382	3.5	2	0.4	5	10.3599	87.2722
7	0.62	553.2656	356.1012	148.7074	3.5	2	0.4	5	10.3636	87.1501
8	0.62	553.5492	355.8328	148.9047	3.5	2	0.4	5	10.3567	87.3401
9	0.62	553.7162	355.5387	148.9067	3.5	2	0.4	5	10.3512	87.5282
10	0.62	553.7973	355.7513	148.9011	3.5	2	0.4	5	10.3522	87.4205
11	0.62	553.2720	355.9848	148.6987	3.5	2	0.4	5	10.3624	87.2162
12	0.62	553.8967	355.3609	148.8487	3.5	2	0.4	5	10.3468	87.6531
13	0.62	553.7864	355.6939	148.9167	3.5	2	0.4	5	10.3517	87.4511
14	0.62	553.8741	355.4008	148.8411	3.5	2	0.4	5	10.3475	87.6275
15	0.62	553.7523	355.6098	148.9153	3.5	2	0.4	5	10.3514	87.4934
16	0.62	553.7623	355.4381	148.8549	3.5	2	0.4	5	10.3495	87.5910
17	0.62	553.5792	355.7680	148.9206	3.5	2	0.4	5	10.3556	87.3806
18	0.62	553.2822	356.0425	148.7034	3.5	2	0.4	5	10.3628	87.1853
19	0.62	553.2584	356.0564	148.7071	3.5	2	0.4	5	10.3633	87.1742
20	0.62	553.6592	355.7637	148.9136	3.5	2	0.4	5	10.3543	87.3942
21	0.62	553.7515	355.6202	148.9149	3.5	2	0.4	5	10.3515	87.4875
22	0.62	553.7875	355.7058	148.9159	3.5	2	0.4	5	10.3518	87.4446
23	0.62	553.7516	355.7842	148.9135	3.5	2	0.4	5	10.3532	87.3957
24	0.62	553.8571	355.4596	148.8446	3.5	2	0.4	5	10.3483	87.5922
25	0.62	553.5675	355.7871	148.9098	3.5	2	0.4	5	10.3560	87.3683
26	0.62	553.7512	355.4762	148.8676	3.5	2	0.4	5	10.3501	87.5681
27	0.62	553.8920	355.4418	148.8497	3.5	2	0.4	5	10.3476	87.6071
28	0.62	553.7520	355.7497	148.9130	3.5	2	0.4	5	10.3528	87.4151
29	0.62	553.3211	355.9303	148.7363	3.5	2	0.4	5	10.3611	87.2536
30	0.62	553.2529	356.0771	148.6975	3.5	2	0.4	5	10.3636	87.1618
31	0.62	553.3767	355.9403	148.7407	3.5	2	0.4	5	10.3604	87.2558
32	0.62	553.6358	355.7797	148.9292	3.5	2	0.4	5	10.3549	87.3820
33	0.62	553.7287	355.4814	148.8411	3.5	2	0.4	5	10.3505	87.5620
34	0.62	553.7241	355.5102	148.8421	3.5	2	0.4	5	10.3508	87.5453
35	0.62	553.5492	355.8245	148.9072	3.5	2	0.4	5	10.3566	87.3448
36	0.62	553.7536	355.4510	148.8541	3.5	2	0.4	5	10.3498	87.5825
37	0.62	553.2988	355.9609	148.7082	3.5	2	0.4	5	10.3617	87.2333
38	0.62	553.3076	355.9518	148.7420	3.5	2	0.4	5	10.3615	87.2397
39	0.62	553.8573	355.4494	148.8423	3.5	2	0.4	5	10.3482	87.5980
40	0.62	553.3144	355.9397	148.7364	3.5	2	0.4	5	10.3613	87.2474
41	0.62	553.2791	356.0428	148.7061	3.5	2	0.4	5	10.3628	87.1847
42	0.62	553.5773	355.8177	148.9011	3.5	2	0.4	5	10.3561	87.3525

**EK 2 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin
MDNSGA-II ile elde edilen Pareto çözüm kümesi (devam)**

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
43	0.62	553.7240	355.4958	148.8422	3.5	2	0.4	5	10.3507	87.5533
44	0.62	553.8801	355.4194	148.8378	3.5	2	0.4	5	10.3476	87.6179
45	0.62	553.2700	355.9591	148.7004	3.5	2	0.4	5	10.3621	87.2303
46	0.62	553.2682	355.9653	148.7018	3.5	2	0.4	5	10.3622	87.2266
47	0.62	553.5810	355.8166	148.9016	3.5	2	0.4	5	10.3561	87.3536
48	0.62	553.7508	355.7617	148.9123	3.5	2	0.4	5	10.3530	87.4082
49	0.62	553.6498	355.7777	148.9317	3.5	2	0.4	5	10.3546	87.3851
50	0.62	553.3847	355.9310	148.7434	3.5	2	0.4	5	10.3601	87.2621

EK 3 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
1	0.98	561.82	416.5	148.8007	3.5	2	0.1	0	4.3008	87.74054
2	0.98	561.8254	416.5	148.6367	3.5	2	0.1	0	4.300881	86.67294
3	0.98	570.334	416.5	157.1977	3.5	2	0.1	0	4.42851	85.40338
4	0.98	569.3828	413.4224	156.2547	3.5	2	0.1	0	4.521959	80.05662
5	0.98	569.3829	413.1391	156.2567	3.5	2	0.1	0	4.531874	79.1326
6	0.98	569.3829	411.9252	156.2567	3.5	2	0.1	0	4.57436	75.57546
7	0.98	566.8608	392.4276	220.5	3.5	2	0.1	0	5.218946	74.90191
8	0.98	566.8608	383.2367	220.5	3.5	2	0.1	0	5.540627	73.5077
9	0.98	608.7829	394.1711	220.5	3.5	2	0.1	0	5.786753	71.56538
10	0.98	608.7829	386.3953	220.5	3.5	2	0.1	0	6.058907	70.57035
11	0.98	665.1286	399.3885	198.8941	3.5	2	0.1	0	6.44933	68.04524
12	0.98	664.1693	390.4233	197.9521	3.5	2	0.1	0	6.748725	64.32861
13	0.98	711.1145	399.5307	220.5	3.5	2	0.1	0	7.134144	62.91224
14	0.98	711.1145	396.4942	220.5	3.5	2	0.1	0	7.240421	61.68309
15	0.98	713.1989	390.0806	220.5	3.5	2	0.1	0	7.496164	56.86416
16	0.98	713.1989	388.8299	220.5	3.5	2	0.1	0	7.539936	53.68674
17	0.98	713.2845	375.0876	220.5	3.5	2	0.1	0	8.022202	51.95568
18	0.98	713.2845	373.2101	220.5	3.5	2	0.1	0	8.087914	51.43763
19	0.98	714.434	358.4653	220.5	3.5	2	0.1	0	8.621226	49.8643
20	0.98	714.434	351.8866	220.5	3.5	2	0.1	0	8.851478	49.25189
21	0.98	776.1969	366.4559	186.4655	3.5	2	0.1	0	9.267997	48.08789
22	0.98	775.2563	354.5522	185.471	3.5	2	0.1	0	9.670517	47.90265
23	0.98	772.5921	346.6801	183.0493	3.5	2	0.1	0	9.906077	46.65003
24	0.98	772.5921	343.1713	183.0493	3.5	2	0.1	0	10.02888	46.57823
25	0.98	774.9868	334.5806	188.4174	3.5	2	0.1	0	10.36548	46.05997
26	0.98	774.9868	332.9352	188.4174	3.5	2	0.1	0	10.42307	45.69385
27	0.98	763.9142	325.5051	183.8059	3.5	2	0.1	0	10.51704	43.53697
28	0.98	763.9142	325.1824	183.8059	3.5	2	0.1	0	10.52833	42.75626
29	0.98	762.0633	318.6999	181.916	3.5	2	0.1	0	10.72745	41.2778
30	0.98	762.0633	317.8674	181.916	3.5	2	0.1	0	10.75659	38.72564
31	0.98	774.0834	318.1733	187.0086	3.5	2	0.1	0	10.92619	35.63661
32	0.98	773.1334	314.9832	186.0202	3.5	2	0.1	0	11.02359	34.17287
33	0.98	772.2896	307.5234	185.0095	3.5	2	0.1	0	11.27202	30.77431
34	0.98	772.2896	305.1951	185.0095	3.5	2	0.1	0	11.35352	30.35657
35	0.98	772.2924	297.4163	184.7492	3.5	2	0.1	0	11.62582	27.29013
36	0.98	772.2924	283.1358	184.7492	3.5	2	0.1	0	12.12563	27.01187
37	0.98	581.14	308.6996	142.1203	7	2	0.1	0	14.38361	25.37118
38	0.98	581.14	303.1753	142.1203	7	2	0.1	0	14.57696	24.69556
39	0.98	582.2069	297.3011	143.3047	7	2	0.1	0	14.79856	21.91008

EK 3 Enerji karma verisinde IRLS modeli için $\hat{Y}_1$ ve $\hat{Y}_2$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	f_1	f_2
40	0.98	583.2046	281.0568	144.2425	7	2	0.1	0	15.38208	20.01363
41	0.98	583.9118	270.0338	144.8262	7	2	0.1	0	15.77849	17.12919
42	0.98	583.9118	265.5617	144.8262	7	2	0.1	0	15.93502	15.39906
43	0.98	728.4981	323.4393	130.9283	7	2	0.1	0	16.07809	13.53496
44	0.98	728.4981	317.1732	130.9283	7	2	0.1	0	16.29741	11.48999
45	0.98	726.9484	313.4322	124.6798	7	2	0.1	0	16.4051	7,10281
46	0.98	726.9484	297.445	124.6798	7	2	0.1	0	16.96465	7.140193
47	0.98	729.1749	294.3178	126.8396	7	2	0.1	0	17.1075	7.077163
48	0.98	729.1749	270.2874	126.8396	7	2	0.1	0	17.94857	6.489888
49	0.98	726.3337	263.2727	124.0968	7	2	0.1	0	18.15146	5.617755
50	0.98	723.337	258.4744	124.0968	7	2	0.1	0	18.3194	5.617199

EK 4 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	f_1	f_2
1	-1	0.6039	0.6771	1.7941
2	-1	-0.4937	0.6073	1.8893
3	-1	-0.0588	0.6517	1.8743
4	-1	0.2292	0.6690	1.8480
5	-1	-0.1262	0.6462	1.8786
6	-1	0.2798	0.6710	1.8420
7	-1	-0.0256	0.6542	1.8719
8	-1	0.2593	0.6700	1.8451
9	-1	-0.4130	0.6172	1.8887
10	-1	0.1891	0.6672	1.8524
11	-1	0.4710	0.6761	1.8157
12	-1	0.3340	0.6729	1.8351
13	-1	-0.0989	0.6485	1.8769
14	-1	-0.4315	0.6150	1.8889
15	-1	-0.1634	0.6430	1.8806
16	-1	0.0772	0.6611	1.8635
17	-1	-0.3052	0.6296	1.8864
18	-1	0.4976	0.6765	1.8116
19	-1	0.5647	0.6770	1.8008
20	-1	0.5963	0.6771	1.7954
21	-1	0.0303	0.6581	1.8676
22	-1	0.3625	0.6738	1.8313
23	-1	0.5355	0.6769	1.8056
24	-1	-0.3845	0.6205	1.8883
25	-1	-0.3797	0.6211	1.8882
26	-1	-0.4650	0.6109	1.8892
27	-1	0.3092	0.6721	1.8383
28	-1	0.4381	0.6755	1.8207
29	-1	-0.3524	0.6241	1.8877
30	-1	0.1201	0.6636	1.8595
31	-1	0.1468	0.6650	1.8568
32	-1	-0.4849	0.6084	1.8892
33	-1	-0.1012	0.6483	1.8771
34	-1	-0.2225	0.6376	1.8834
35	-1	-0.3516	0.6242	1.8877
36	-1	0.4180	0.6751	1.8236
37	-1	-0.3272	0.6269	1.8871
38	-1	0.0543	0.6596	1.8655

EK 4 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	f_1	f_2
39	-1	-0.1806	0.6415	1.8815
40	-1	-0.2024	0.6395	1.8826
41	-1	-0.2628	0.6336	1.8850
42	-1	-0.2294	0.6369	1.8837
43	-1	0.0078	0.6565	1.8694
44	-1	0.1072	0.6628	1.8607
45	-1	0.0549	0.6597	1.8655
46	-1	-0.4530	0.6124	1.8891
47	-1	0.0182	0.6573	1.8686
48	-1	0.5160	0.6767	1.8087
49	-1	0.1625	0.6658	1.8552
50	-1	-0.1991	0.6398	1.8824

EK 5 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	f_1	f_2
1	-1.414	0.2260	0.6176	2.0999
2	-1.414	0.4781	0.6322	2.0385
3	-1.414	0.3150	0.6236	2.0816
4	-1.414	0.1176	0.6090	2.1171
5	-1	-0.0532	0.6521	2.0295
6	-1	0.5498	0.6770	1.9121
7	-1	0.2081	0.6681	1.9996
8	-1.414	0.3962	0.6283	2.0617
9	-1	0.3138	0.6723	1.9784
10	-1	-0.1569	0.6436	2.0325
11	-1	0.1200	0.6636	2.0133
12	-1.414	-0.1563	0.5813	2.136
13	-1.414	0.3133	0.6235	2.082
14	-1	0.3222	0.6725	1.9765
15	-1	-0.1189	0.6469	2.032
16	-1	0.4758	0.6762	1.9357
17	-1.414	-0.0664	0.5913	2.1336
18	-1.414	0.4105	0.6290	2.0579
19	-1.414	0.1213	0.6093	2.1166
20	-1.414	0.4965	0.6330	2.0328
21	-1	-0.0800	0.6500	2.0307
22	-1.414	0.2365	0.6183	2.0979
23	-1	0.1225	0.6637	2.0129
24	-1	0.2060	0.6680	2
25	-1	0.4814	0.6763	1.934
26	-1	-0.0823	0.6498	2.0308
27	-1	0.2034	0.6678	2.0004
28	-1.414	0.4091	0.6289	2.0583
29	-1.414	0.4734	0.6320	2.0399
30	-1.414	-0.0664	0.5913	2.1336
31	-1	-0.1481	0.6444	2.0324
32	-1.414	0.2343	0.6182	2.0983
33	-1.414	0.3181	0.6238	2.0809
34	-1	0.1170	0.6634	2.0137
35	-1	0.3135	0.6722	1.9785
36	-1.414	0.1009	0.6075	2.1192
37	-1.414	0.4303	0.6300	2.0524
38	-1	0.1157	0.6633	2.0138
39	-1	0.3089	0.6721	1.9795
40	-1.414	0.3079	0.6231	2.0832

EK 5 Gıda karma verisi için $\hat{Y}_2$ ve $\hat{Y}_3$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	f_1	f_2
41	-1	0.2033	0.6678	2.0004
42	-1	0.4254	0.6753	1.9503
43	-1.414	0.4923	0.6328	2.0341
44	-1	-0.0756	0.6504	2.0306
45	-1.414	0.1169	0.6089	2.1172
46	-1	-0.1362	0.6454	2.0323
47	-1.414	-0.0596	0.5920	2.1333
48	-1.414	0.2082	0.6162	2.1031
49	-1	-0.0757	0.6504	2.0306
50	-1.414	0.4105	0.6290	2.0579

EK 6 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	f_1	f_2
1	-1.414	-1.414	1.7699	0.0136
2	-1	-0.4936	1.8893	0.26
3	-1	-0.9128	1.8754	0.1946
4	-1	-0.9746	1.871	0.185
5	-1	-1.0159	1.8677	0.1785
6	-1.414	-1.3752	1.7755	0.021
7	-1.414	-1.2307	1.7939	0.0485
8	-1	-0.8423	1.8797	0.2056
9	-1.414	-1.1810	1.7995	0.058
10	-1	-1.2648	1.8423	0.1397
11	-1.414	-1.3508	1.7788	0.0256
12	-1	-0.7488	1.8841	0.2202
13	-1.414	-0.917	1.8227	0.1082
14	-1	-1.3902	1.8258	0.1201
15	-1	-0.6762	1.8866	0.2315
16	-1	-1.0684	1.8632	0.1703
17	-1	-1.3289	1.8341	0.1297
18	-1	-1.2234	1.8472	0.1461
19	-1	-1.2895	1.8392	0.1358
20	-1	-1.235	1.8458	0.1443
21	-1.414	-1.2497	1.7917	0.0449
22	-1	-0.7217	1.8851	0.2244
23	-1.414	-0.9920	1.8172	0.0939
24	-1.414	-1.1057	1.8073	0.0723
25	-1	-0.6368	1.8876	0.2377
26	-1	-1.0838	1.8617	0.1679
27	-1.414	-1.0675	1.8108	0.0796
28	-1.414	-1.015	1.8154	0.0896
29	-1.414	-0.9288	1.8219	0.106
30	-1	-1.1552	1.8547	0.1568
31	-1.414	-1.3317	1.7814	0.0293
32	-1	-0.8026	1.8817	0.2118
33	-1	-1.1224	1.8580	0.1619
34	-1	-1.1913	1.8508	0.1512
35	-1	-1.1124	1.8590	0.1635
36	-1.414	-1.16	1.8018	0.0619
37	-1	-1.1593	1.8543	0.1562
38	-1.414	-0.9601	1.8197	0.1
39	-1.414	-1.2686	1.7894	0.0413
40	-1	-1.1885	1.8511	0.1516
41	-1.414	-1.1081	1.807	0.0718
42	-1	-1.3727	1.8282	0.1229
43	-1	-0.5573	1.8889	0.2501

EK 6 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-II ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	f_1	f_2
44	-1.414	-1.3217	1.7827	0.0312
45	-1	-0.8125	1.8812	0.2103
46	-1.414	-1.1335	1.8045	0.067
47	-1.414	-1.0325	1.8139	0.0862
48	-1	-1.2889	1.8393	0.1359
49	-1.414	-1.1375	1.8041	0.0662
50	-1.414	-1.309	1.7843	0.0336

EK 7 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi

N	x_1	x_2	f_1	f_2
1	-1.414	-1.0204	1.9644	0.0885
2	-1.414	-1.1931	1.8880	0.0556
3	-1.414	-1.3485	1.8074	0.0261
4	-1.414	-1.3945	1.7813	0.0173
5	-1.414	-1.2838	1.8423	0.0384
6	-1.414	-0.8496	2.0261	0.1210
7	-1.414	-1.3753	1.7923	0.0210
8	-1.414	-1.1217	1.9213	0.0692
9	-1.414	-1.4302	1.7604	0.0105
10	-1.414	-1.2312	1.8693	0.0484
11	-1.414	-1.4144	1.7697	0.0135
12	-1.414	-1.3244	1.8206	0.0306
13	-1.414	-0.1757	2.1360	0.2493
14	-1.414	-1.1960	1.8866	0.0551
15	-1.414	-1.2709	1.8491	0.0408
16	-1.414	-1.4076	1.7737	0.0148
17	-1.414	-1.3996	1.7784	0.0163
18	-1.414	-1.3563	1.8030	0.0246
19	-1.414	-0.8651	2.0211	0.1181
20	-1.414	-1.3792	1.7901	0.0202
21	-1.414	-0.9936	1.9750	0.0936
22	-1.414	-0.2647	2.1337	0.2324
23	-1.414	-1.2312	1.8693	0.0484
24	-1.414	-1.4295	1.7608	0.0106
25	-1.414	-1.1236	1.9205	0.0689
26	-1.414	-1.3150	1.8257	0.0324
27	-1.414	-1.2311	1.8693	0.0484
28	-1.414	-1.0154	1.9664	0.0895
29	-1.414	-1.3730	1.7936	0.0214
30	-1.414	-1.1943	1.8875	0.0554
31	-1.414	-1.4239	1.7641	0.0117
32	-1.414	-1.1138	1.9248	0.0707
33	-1.414	-1.3587	1.8017	0.0241
34	-1.414	-1.2777	1.8455	0.0395
35	-1.414	-1.3939	1.7816	0.0174
36	-1.414	-0.1952	2.1358	0.2456
37	-1.414	-1.4156	1.7690	0.0133
38	-1.414	-1.3122	1.8272	0.0330
39	-1.414	-0.8496	2.0261	0.1210
40	-1.414	-1.4293	1.7609	0.0107

EK 7 Gıda karma verisi için $\hat{Y}_3$ ve $\hat{Y}_4$ yanıt fonksiyonlarına ilişkin MDNSGA-III ile elde edilen Pareto çözüm kümesi (devam)

N	x_1	x_2	f_1	f_2
41	-1.414	-1.0092	1.9689	0.0907
42	-1.414	-0.2390	2.1347	0.2373
43	-1.414	-1.2806	1.8440	0.0390
44	-1.414	-1.2311	1.8693	0.0484
45	-1.414	-0.7524	2.0552	0.1395
46	-1.414	-1.3739	1.7931	0.0212
47	-1.414	-1.1921	1.8885	0.0558
48	-1.414	-1.3563	1.8030	0.0246
49	-1.414	-1.3163	1.8250	0.0322
50	-1.414	-1.1236	1.9205	0.0689