

OPTİMİZASYON TEKNİKLERİ-5 Hafta

GENETİK ALGORİTMA

1. Giriş

Geniş çözüm uzaylarının klasik yöntemlerle taranması hesaplama zamanını artırmaktadır. Genetik algoritma ile kabul edilebilir doğrulukta kısa sürede bir sonuca gidilebilir. Genetik algoritmalar özellikle çözüm uzayının geniş, süreksiz, karmaşık olduğu durumlarda başarılı sonuçlar vermektedir.

Genetik algoritma bir arama ve optimizasyon tekniğidir. Genetik algoritma tek bir noktadan değil noktalar kümesinden geniş bir alandan arama yapmaktadır. Olasılık içeren kurallar kullanır. Türev ve benzeri yardımcı bilgileri kullanmaz. Yalnızca amaç fonksiyonu bilgisine ihtiyaç duyar.

Genetik algoritma genel anlamda, dizilerden oluşan bir popülasyona (nesil) çoğalma, çaprazlama ve mutasyon operatörlerinin uygulanmasını içerir. Bu operatörlerin uygulanmasından sonra yeni bir popülasyon (yavru nesil) oluşur. Yeni nesil eski ebeveyn nesil ile yer değiştirir. Her dizinin (kromozomun) bir uyum değeri vardır. Yeni nesiller (yeni diziler) bu uyum değerlerin göre seçilirler. Her yeni üretilen nesilde daha uyumlu iyi nesiller üretilmeye çalışılır.

2. Temel Bilgiler

Genetik Algoritmanın Uygulanması

Uygulamadaki ilk adım, ilk popülasyonun oluşturulup, uyum değerinin hesaplanmasıdır. Daha sonra mevcut nesile, temel genetik operatörler (çoğalma, çaprazlama, mutasyon) uygulanır. Her nesil için uyum değeri hesaplanır. Bu durum durdurma kriteri sağlanana kadar devam eder. Adımlar şu şekilde açıklanabilir.

- Arama uzayındaki tüm olası çözümlerden bir grup çözüm dizi olarak kodlanır. Genellikle rastsal bir işlem yapılır ve başlangıç popülasyonu oluşturulur.
- Her bir dizi için uyum değeri hesaplanır. Bulunan uyum değerleri dizilerin çözüm kalitesini gösterir.
- Bir grup dizi belirli bir olasılık değerine göre rastsal olarak seçilir. Seçilen diziler çaprazlama ve mutasyon işlemlerine tabi tutulur.
- Oluşan yeni popülasyon eski popülasyon ile yer değiştirilir.
- Durdurma kriteri sağlanana dek yukarıdaki işlemlere devam edilir. En uygun olan dizi çözüm olarak seçilir.

Kodlama

Çözümlerin kodlanması, probleme özgü bilgilerin genetik algoritmanın kullanacağı şekle çevrilmesidir. Her problem kendine özgü farklı kodlamalara ihtiyaç duyabilir. Tüm problemler için geçerli olabilecek genel bir kodlama tekniği yoktur. Kodlama türleri 3 grupta incelenebilir.

İkili Kodlama: En yaygın kodlama yöntemidir. Problemin olası çözümleri 0 ve 1 sayıları ile kodlanır.

Sıralı Kodlama: Genellikle birleşik optimizasyon problemlerinde kullanılmaktadır. Dizinin uyum değeri Gen değerlerine ve genlerin sıralarına da bağlıdır. Gezgin satıcı problemi ve Araç rotalama problemleri bu kodlamaya uygundur.

Gezgin satıcıda her tamsayı değeri bir şehri gösterir. Bu yöntem ile tekrar eden değerler engellenmiş olur. Ayrıca uyum değerlerinin hesaplanması da basit hale gelmiş olur. Örneğin $X=[4 \ 3 \ 0 \ 1 \ 2]$ dizisi, sırayla dördüncü, üçüncü, merkez şehrin, birinci ve ikinci şehirlerin ziyaret edileceğini anlatmaktadır.

Değer Kodlaması: Değerler kodlamasında her dizi, bir değerler kümesinden oluşmaktadır. Değerler probleme göre herhangi bir sayı yada karakterler olabilir. Örnek:

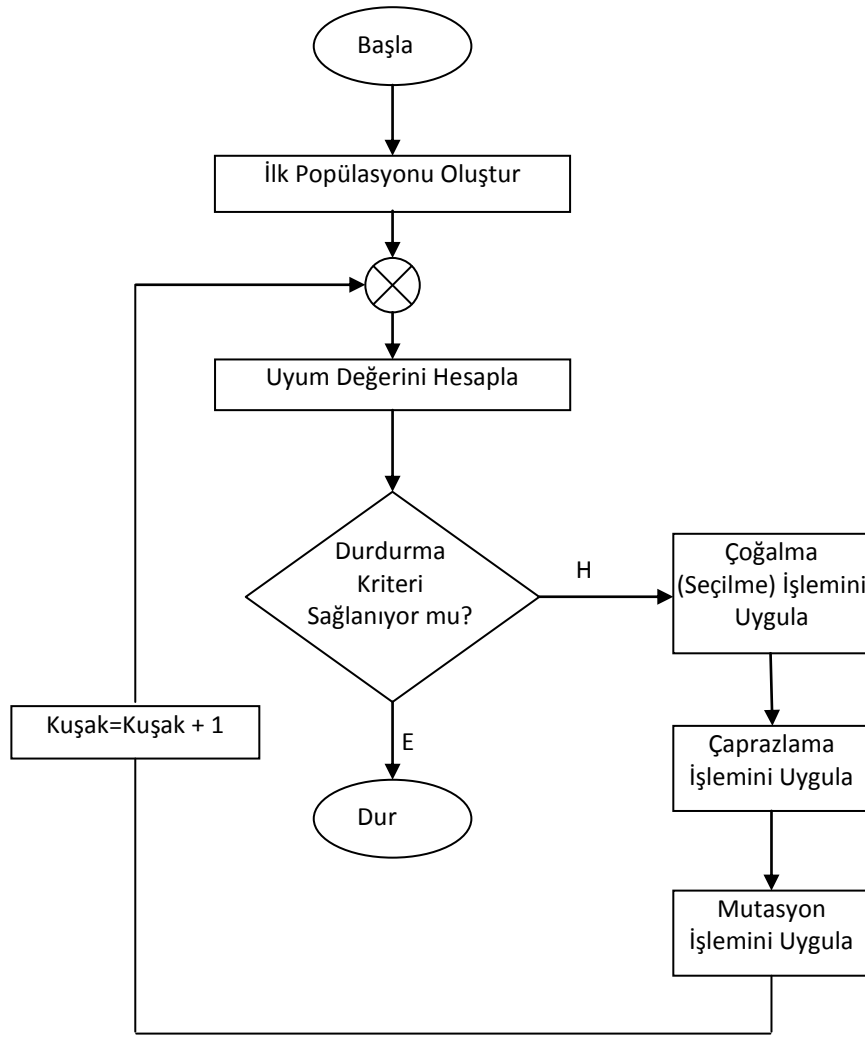
Dizi A= 1.2324, 5.3243, 0.4556, 2.3293, 2.4545

Dizi B= ABDJEIFRGT

Dizi C= (geri), (geri), (sağ), (ileri), (sol)

Değer kodlaması bazı özel problemler için oldukça kullanışlıdır. Fakat bu tip kodlama kullanıldığında probleme özgü çaprazlama ve mutasyon yöntemleri gerekir. Böyle bir kodlama yöntemi Yapay Sinir Ağlarında ağırlıkların bulunması amacıyla kullanılabilir. Bu durumda dizide kullanılan gerçek sayılar, yapay sinir ağlarındaki ağırlıkları ifade eder.

Ağaç Kodlama Yöntemi: Bu yöntemde her dizi nesnelere oluşan bir ağaç yapısı şeklinde ifade edilir. Bu nesnelere fonksiyonlar yada programlama dillerinde geçen komutlar olabilir. Bu yöntem verilen değerlere uygun bir fonksiyon bulmak için kullanılabilir.



Şekil 1. Genetik algoritmanın akış diyağramı.

Uyum Değeri

Dizi ile taşınan genetik bilgi her kuşakta uyum değeri daha iyi dizilerin seçilmesine rehberlik eder. Dizinin uyum değeri ne kadar yüksek ise, yaşama ve çoğalma şansında o kadar yüksek olacaktır.

Problemin türünü göre Uyum değeri farklı şekillerde olur. Amaç fonksiyonu kârı en yüksek tutmak yada, maliyetleri en düşük tutmak olabilir. Eğer amaç kâr ise bu durumda problem şöyle olabilir.

“ $f(x)$ maksimum yapan x değerini bulun”

Amaç maliyeti en düşük tutmak yada Hatayı minimum yapmak $E(x)$ ise problem şu şekilde olur.

“ $E(x)$ 'i minimum yapan x değerini bulun”

Alt Dizi Kavramı

Alt diziler genetik algoritmaların davranışlarını açıklamak için kullanılan bir yapıdır. Bir alt dizi belirli dizi kümeleri arasındaki benzerliği tanımlayan bir dizidir. Alt diziler $\{0,1,*\}$ harfleri kullanılarak tanımlanır.

Örneğin $H=01*1*$ alt dizisi, ilk konumunda 0, ikinci ve dördüncü konumunda 1 değeri olan diziler kümesi içindir. * sembolü dizinin o konumunun hangi değeri alıp almadığının önemli olmadığını gösterir. Dizi bu konumlarda 0 yada 1 de olabilir. Fakat önemli değildir.

Eğer bir dizi alt dizinin örüntüsüne uyarsa bu diziyeye “H nin bir örneğidir” denir. Alt dizilerin iki özelliği mevcuttur. Bu özellikler aşağıda verilmiştir.

Alt dizi derecesi: Alt dizi örüntüsünde bulunan sabit konumların sayısıdır. Örneğimizde 0 ve 1 değerlerinin sayısının toplamına eşittir. $o(H)$ ile gösterilir.

Alt dizi uzunluğu: Alt dizi örüntüsündeki ilk ve son konumlar arasındaki uzaklığı gösterir. $\delta(H)$ ile gösterilir.

Alt dizi derecesi ve uzunluğu genetik algoritmaların temelinde önemli bir yer tutar. Derecesi düşük, uzunluğu kısa olan alt diziler “yapı blokları” olarak adlandırılır. Genetik algoritmaların işleyişinde uygun yapı bloklarının tanımlanmasını ve yapı bloklarının daha uygun yapı blokları elde etmek için birleştirilmesi tavsiye edilir.

“Popülasyon ortalamasının üstünde uyum gücü gösteren, düşük dereceye ve kısa uzunluğa sahip alt diziler zamanla üstsel olarak çoğalırlar. Bu çoğalma sonucu ebeveyn dizilerden daha üstün özelliklere sahip diziler ortaya çıkmaktadır.

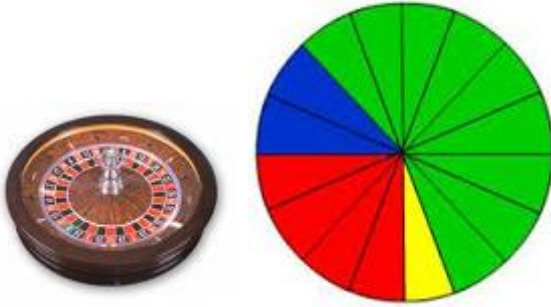
Genetik Algoritma Operatörleri

Genetik algoritmada kullanılan temel operatörler (işlemler) Çoğalma operatörü (Seçilme Operatörü), Çaprazlama operatörü ve Mutasyon Operatörüdür.

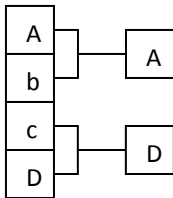
Çoğalma Operatörü (Reproduction) (Seçilme Operatörü)

Çoğalma işlemi, dizilerin uyum fonksiyonu (amaç fonksiyonu) değerlerine göre kopyalandığı süreçtir. Yüksek uyum değerine sahip dizilerin bir sonraki kuşağa katkılarının daha yüksek olacağı anlamına gelir. Çoğalma işlemi, dizileri seçme işleminden, seçilmiş dizileri bir eşleme havuzuna kopyalama işleminden, havuzda dizileri genelde çiftler halinde gruplara ayırma işleminden oluşmaktadır. Literatürde bu operatör, seçilme operatörü olarak da geçmektedir. Seçim işlemi bir sonraki kuşak için yavru üretmek amacıyla hangi ailelerin yer alması gerektiğine karar verir. Bu amaçla ortalama uyum değeri üzerindeki dizilere çoğalma fırsatı tanıyarak olur. Literatürde çoğalma yöntemleri Rulet tekerleği (roulette Wheel), turnuva (tournament), sabit durum (steady state) ve sıralama (permutation) yöntemleri sayılabilir.

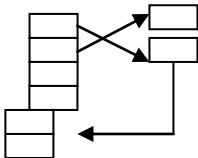
Rulet Tekerleği Yöntemi: Bu yöntemde her dizi uyum değeriyle orantılı bir olasılık değeri ile seçilmektedir. Rulet tekerleğinin yüzeyi dizilerin uyum değerleri ile orantılı olarak işaretlenir. Tekerlek kaç defa döndürülürse her seferinde bir dizi eşleme havuzuna atılır. Daha iyi uyum değerine sahip diziler tekerlekte daha fazla yer aldıklarından onların seçilme şansları daha yüksektir.



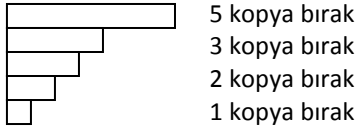
Turnuva Yöntemi: Tüm nesil içinden her seferinde rastgele iki dizi seçilir. Bunlar turnuvaya katılmış olur. Bu ikisi arasından yapılan turnuvada hangisi daha iyiye o kazanır. Her kişinin (dizinin) turnuvaya katılması iki kez yapılır. Dolayısıyla en iyi kişi (dizi) her iki turnuvada başarılı olur. En kötü kişi ise her iki turnuva da başarısız olur. Her iki turnuvada başarılı olan listeye ismini yazdırır. Böylece en başarılı olanlar iki kez orta seviyede olanlar ise bir kez isimlerini yazdırabilirler.



Sabit Durum Yöntemi: Bu yöntemde popülasyon içindeki sınırlı sayıda dizinin yer değiştirmesini sağlamak, bir dizinin birden fazla kuşak hayatta kalmasını sağlamak için kullanılır. Bu amaçla mevcut popülasyon içerisinde belli sayıda dizi seçilir. Bunlar yeni nesiller üretir. Üretilen yeni nesiller tekrar popülasyona katılır. Fakat popülasyon nüfusunun artmaması için içeriden bazılarının atılması gerekir. Bu nedenle burada karar verilmesi gereken iki işlemden bahsedilebilir. Yeni nesil üretecek dizilerin hangi kriterle seçileceği, Yeni doğan nesillerin ise kimin yerini alacağına nasıl karar verileceğidir.



Sıralama Yöntemi: Bu yöntemde tüm diziler uyum değerlerine göre artan bir şekilde sıralanır. Uygun bir fonksiyonla en yüksek uyumlular havuza daha fazla kopya bırakır, en kötü diziler ise kopya elde edememiş olur. Bu yöntemde arama hızı yüksektir.



Çaprazlama Operatörü(Crossover)

Mevcut gen havuzunun potansiyelini artırmak üzere, bir önceki kuşaktan daha iyi nitelikleri yeni diziler oluşturmak için çaprazlama operatörü kullanılır. Çoğalma işlemi sonucunda elde edilen yeni populasyondan rastgele iki dizi seçilmekte ve karşılıklı çaprazlama işlemine tabi tutulmaktadır.

Çaprazlama işlemi seçilen iki dizinin ortasından rastgele alınan bir noktadan karşılıklı olarak çaprazlama yapılarak gerçekleştirilir. Uygun bir çaprazlama noktası seçildiğinde ebeveynlerden uyum değeri daha yüksek diziler elde edilebilmektedir. Fakat uygun çaprazlama noktası genellikle bilinmediğinden rastsal bir nokta seçilmektedir.

Temel çaprazlama yöntemleri tek noktalı ve çift noktalı çaprazlama yöntemleridir.

Tek noktalı çaprazlama yöntemi: Havuzdaki diziler rastgele eşleşirler. Seçilen her dizi çifti için ilk ve son gen dışında aradaki genlerden rastgele bir yer seçilir. Burası çaprazlama noktasını gösterir. Bu noktadan sonra gelen genler her iki dizide karşılıklı olarak yer değiştirir. Bu işlem için diziler aynı uzunlukta olmalıdır.

Aşağıdaki dizilerde 3. konum çaprazlama noktası olsun. Bu noktadan sonraki genler çaprazlandığında yeni diziler elde edilir.

$$P1 = 1,0,0,[1,1,1,1,0] \Rightarrow P1' = 1,0,0,[1,0,0,1,0]$$

$$P2 = 1,0,1,[1,0,0,1,0] \Rightarrow P2' = 1,0,1,[1,1,1,1,0]$$

Çift noktalı çaprazlama yöntemi: Bu yöntemde dizi üzerinde ilk ve son genler hariç iki tane rastgele nokta seçilir. Çaprazlama işlemi seçilen bu iki nokta arasındaki genlerin yer değişimidir. Ve bu noktalar arasında genler değiştirilir.

Aşağıdaki dizilerde 1. ve 5. genden sonra gelen aralık çaprazlama noktalarımız olsun. Bu iki nokta arasında kalan genleri çaprazlayalım.

$$P1 = 1,[0,0,1,1],1,1,0 \Rightarrow P1' = 1,[0,1,1,0],1,1,0$$

$$P2 = 1,[0,1,1,0],0,1,0 \Rightarrow P2' = 1,[0,0,1,1],0,1,0$$

PMX Çaprazlama Yöntemi (Sıra temelli): Bu yöntem gezgin satıcı ve araç rotalama problemlerinde yaygın olarak kullanılmaktadır. Çünkü bu problemlerde dizileri oluşturan genlerin aynı dizi içinde tekrar etmemesi gerekir. Bu yöntem bu tekrarı engelleyen bir yöntemdir. Ayrıca bu yöntemde seçilen bir alt dizinin korunması sağlanmaktadır. Bu alt dizi çocuk diziyeye aktarılmaktadır. Bu yöntemde iki ebeveyn diziden yalnız bir adet yavru dizi elde edilmektedir.

Yöntem şu şekildedir.

Havuzdan iki ebeveyn dizi rastgele seçilsin. Bunlar P1 ve P2 olsun.

$$P1=1,2,4,6,3,7,5,8 \Rightarrow \text{dizisinden } 4,6,3 \text{ nolu alt dizisi rastgele seçilsin.}$$

$$P2=5,4,1,7,2,6,8,3$$

$$C= _ _ 4,6,3, _ _ _ \text{ yavru dizi olur.}$$

İkinci diziden bu sayıların olduğu yerleri çıkaralım.

$P2= 5,*,1,7,2,*,8,*$ olur. Bu dizi içerisindeki 1,7,2 sayıları 4,6,3 alt dizisine karşılık gelen konumdadır. 1,7,2 dizisi * ile gösterilen konumlara Yavru dizi içerisine yerleştirilelim. Bu durumda yavru dizi şöyle olur.

$C= _ ,1,4,6,3,7, _ ,2$ olur. İkinci dizideki 5 ve 8 sayıları aynı konumlarına çocuk dizide yerine yazılırsa en son çocuk dizi şu hali alır.

$$C= 5,1,4,6,3,7,8,2$$

Böylece iki ebeveyn diziden bir çocuk dizi üretilmiş olur.

OX Çaprazlama (Sıra temelli): bu yöntemde yine dizideki genlerin tekrarlı olmasını engellemektedir. Sıra temelli yapıya sahiptir. PMX den farklı olarak burada iki ebeveyn diziden yine 2 tane çocuk dizi oluşturulmaktadır. PMX de olduğu gibi yine ebeveyn dizilerden seçilen alt diziler korunarak çocuk dizilere aktarılmaktadır.

Yöntem şu şekildedir. İki tane ebeveyn dizi alalım.

P1= 1,3,5,**7,9,2**,4,6,8 => dizisinden [7,9,2] alt dizisini seçelim. Bu dizinin aynı konumuna ikinci diziden karşılık gelen alt diziyide bulalım.

P2= 7,5,9,**2,6,3**,1,4,8 => [2,6,3] dizisi yukarıdaki alt diziyile aynı konuma karşılık gelir.

Bu iki alt dizi çocuk dizilere aynı konumda olacak şekilde aktarılırlar.

C1= __,__,7,9,2,__

C2= __,__,2,6,3,__ çocuk diziler bu hale gelir.

Ebeveyn diziler, ikinci kesim noktasından (alt diziden sonra başlayan kısım) başlayarak dizi tekrar bir daha sıralanır.

P1'=4,6,8,1,3,5,**7,9,2**

P2'=1,4,8,7,5,9,**2,6,3**

Yeni sıralanan birinci ebeveynden ikincinin alt dizisi çıkarılır. Aynı şekilde ikinci ebeveynden de birincinin alt dizisi çıkarılır.

P1''=4,* ,8,1,* ,5,7,9,*

P2''=1,4,8,* ,5,* ,* ,6,3

Çıkarılan bu konumlar olmadan dizi düzenlenir.

P1'''=**4,8,1,5,7,9**

P2'''=**1,4,8,5,6,3**

Elde edilen bu yeni kısa diziler P1''' olan C2 çocuk dizisinin ikinci kesim noktasından sonra (alt dizi bitiminden) tekrar sırayla yerleştirilir. (Dikkat ilk 3 eleman yerleştirildiğinden dizi sonuna gelince yerleştirme baştan devam edilir.). Aynı şekilde P2''' elemanları da C1 çocuk dizisine yerleştirilir.

C1= **5,6,3**,7,9,2,**1,4,8**

C2= **5,7,9**,2,6,3,**4,8,1**

Böylece içerisinde tekrarlı sayıların bulunmadığı iki yeni çocuk dizi elde edilmiş olur.

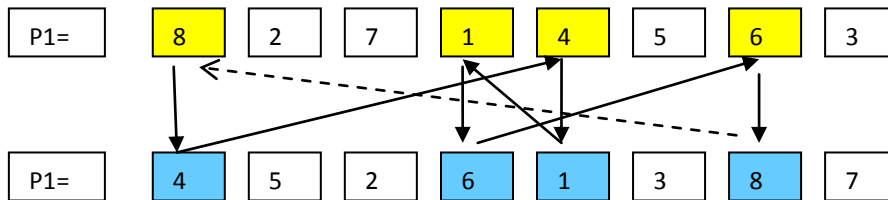
CX Çaprazlama (Cycle Crossover)(Sıra temelli): Bu yöntemde diğer iki yöntem gibi sıra temelli (dizide tekrarlı gen yoktur) yöntemdir. Bu yöntemde her seçilen ebeveyn çiftinden iki tane yavru dizi üretilir. Ebeveyn dizilerde çaprazlama noktası seçilmez. Yöntem şu şekildedir.

Ebeveyn dizilerimiz şu şekilde olsun.

P1 =**8,2,7,1,4,5,6,3**

P2 =**4,5,6,2,1,3,8,7**

İlk ebeveyn dizinin en sol konumundan işleme başlanır. Her bir diziden seçilen genler çocuk dizilere yerleştirilir. İlk gen 8 dir. Bu sayı ile aynı konumda diğer dizide 4 vardır. 4 sayısı birinci dizide beşinci konumdadır. Aynı konumda ikinci dizide 1 vardır. 1 sayısı birinci dizide dördüncü konumdadır. Aynı konumda ikinci dizide 6 vardır. 6 sayısı birinci dizide yedinci konumdadır. Aynı konumda ikinci dizide 8 vardır. Fakat 8 sayısı daha önce yerleştirildiği için birinci diziden seçim yapılamaz. Bulunan bu ilk genler çocuk dizilere yerleştirilir.



Buna göre çocuk diziler şu şekilde olur.

C1 =**8,* ,* ,1,4,* ,6,***

C2 =**4,* ,* ,6,1,* ,8,***

Kalan * yıldızlı kısımlar ise her iki dizide karşılıklı olarak değiştirilir.

C1 =**8,5,2,1,4,3,6,7**

C2 =**4,2,7,6,1,5,8,3**

Böylece iki adet ebeveynden yeni iki adet çocuk dizi üretilmiş olur.

LOX Çaprazlama (Linear Order Crossover): İş çizelgeleme problemlerinde kullanılan bir yöntemdir. Diziler iki noktadan bölgelere ayrılır. Uygun bir şekilde çaprazlanarak yeni nesiller üretilir.

Ağaç Çaprazlama: Alt kısımları bulunan nesnel yapı şeklindeki dizileri çaprazlarken kullanılır. Bir ağacın (dizinin) kesilen alt dalları başka bir ağacın dallarına monte edilir.

Mutasyon Operatörü (Mutation)

Çaprazlama işlemi, mevcut gen potansiyelini araştırır. Eğer populasyon problemi çözmek için gereksinim duyulan tüm şifrelenmiş bilgiyi içermezse tatmin edici bir çözüme ulaşamaz. Bu nedenle mevcut gen havuzundan yeni diziler üretme yeteneğine sahip bir operatöre ihtiyaç duyulur. Bu görevi mutasyon gerçekleştirir. Böylece genetik çeşitlilik artırılmış olur. Mutasyon problemin herhangi bir yerel optimum noktaya yakınsamasını engelleyebilmektedir. Mutasyon oranları genellikle düşük tutulmaktadır. Bunun nedeni çaprazlama sonucu elde edilen uyum değeri yüksek dizileri kaybetmemektir.

Değer Değiştirme Yöntemi: Basit genetik algoritmada dizideki herhangi bir konumun değerinin 1 iken 0 yada tersini yaparak gerçekleştirilir.

P1= 1,0,0,1,1,1,1,0

P1'=1,0,0,1,1,0,1,0

Kayıdırma Yöntemi: Dizi içerisinde rastsal olarak belirlenen bir blok genin yine aynı dizi içerisinde rastgele bir konuma yerleştirilmesidir.

P1 = AKDE**EM**ZIF

P1'= A**EM**ZIKDF

Yerleştirme Yöntemi: Dizi içinde rastgele bir bir genin seçilerek, yine aynı dizi içinde rastgele başka bir konuma yerleştirilmesidir.

P1 = AK**D**EMZIF

P1'= ADEM**Z**KIF

Karşılıklı Değişim Yöntemi: Dizi içerisinde Rastsal olarak iki gen seçilmekte ve bu iki genler birbiriyle yer değiştirmektedir.

P1 = AK**D**EMZIF

P1'= A**Z**DEMKIF

Durdurma Kriteri

Çoğalma, çaprazlama ve mutasyon işlemlerinden sonra yeni bir kuşak oluşmaktadır. Tüm bu işlemler sonsuz döngü içerisinde yapılır. Eğer bir durdurma kriteri belirlenmez ise bu süreç sonsuza kadar devam eder. Bununla ilgili olarak şu yöntemler kullanılmaktadır.

Hesaplama Zamanı Kriteri: Bu yöntemde önceden bir hesaplama zamanı veya döngü sayısı belirlenmekte, bu zaman veya döngü sayısına ulaşıldığında durdurulmaktadır. Bu yöntemin sakıncaları olabilmekte. Döngü devam ettirildiğinde iyileşmeler görülebilecektir. Belirlenen döngü sayısı gerektiğinden de fazla olabilir.

Optimizasyon Hedefi Kriteri: Önceden ulaşılması istenen amaç fonksiyonu değeri bilinmektedir. Uyum değeri bu değere ulaştığında algoritma durdurulmaktadır.

Minimum iyileşme kriteri: Genetik algoritma problemlerinde bulunan en iyi çözümler önce hızlı daha sonra yavaş yavaş artış göstermektedir. Bulunan değerlerdeki iyileşme hızının giderek azalması ve sifıra yaklaşması, artık daha fazla iyileşme beklenmemesi gerektiğini gösterebilir. Çözüme harcanacak zaman ile çözümden beklenecek kalite arasında bir denge kurularak durdurulur.