

# 1. Genetikus algoritmusok elméleti alapjai. Optimalizálási feladat, egyed, fenotípus és genotípus alak, populáció, bináris és real genetikus algoritmus. Egyszerű bináris genetikus algoritmus (SGA) blokkvázlata.

**Koncepció:** Egyelőre csak egyetlen célfüggvény optimalizálásával foglalkozunk.

**Feladat:** A globális optimum meghatározása (noha több lokális minimum lehetséges), ami minimumproblémává fogalmazható:

$$\min f(x_1, \dots, x_{N_{var}})$$

## Jellemzők/Feltételek:

- Ismert minden  $x_i$  változó véges legkisebb és legnagyobb értéke.
- $x_i$  típusa tipikusan valós ( $x_i \in R^1$ ), esetleg egész ( $x_i \in Z^1$ ).
- $x_i$  ábrázolása
  - o valós: lebegőpontos számként vagy előírt pontosságig (2-es kompl., Gray-kód)
  - o egész: alkalmas bázissal, mely integer számmal fejezhető ki
- $x_1, \dots, x_{N_{var}}$  kombinációi az egyedek
  - o Real alak: fenotípus (= az egyed érzékelhető, megállapítható - leírható és megmérhető - külső és belső tulajdonságainak összessége, az örökletes alap és az életfeltételek kölcsönhatásának eredője)
  - o Kódolt alak: genotípus/kromoszóma (= a szervezet kromoszómális génjeiben tárolt genetikai információk összessége, amely meghatározza a szervezet külső megjelenését/fenotípusát)
  - o Az  $x_i$  változónak megfelelő kombinációrészlet helyén áll a gén konkrét allélje (= valamely gén szerkezetileg és funkcionálisan eltérő módosulatainak egyike)
  - o A közvetlenül  $x_1, \dots, x_{N_{var}}$  real kombinációként ábrázolt egyedet is kromoszómaként tekintjük (ekkor genotípus = fenotípus)
- A populáció  $N_{ind}$  darab egyedből áll.
- A rekombinációhoz (= keresztezés: kicserélődés eltérő genotípust meghatározó genetikai anyagok között, amelynek eredménye különbözik a szülők genotípusától) kiszelektált egyedek száma:  $N_{ind} \cdot GGAP$ , ahol GGAP a generation gap.
- Az optimum helyének meghatározása globális sztochasztikus kereső eljárással:
  - o Elv: szelekció, rekombináció, visszahelyezés, mutáció és migráció (+ elitista stratégia)
  - o Lépések realizálása: sztochasztikus (véletlenszámra alapozott) eljárások
- Populációk száma szerint:
  - o SGA (Single Genetic Algorithm)
  - o MPGA (Multi Population Genetic Algorithm)
- Implementálás: MATLAB
- Célfüggvények legyenek egységes felépítésűek és cserélhetők
- Célfüggvény kiszámításához: genotípus (kromoszóma)  $\rightarrow$  fenotípus konverzió (bs2rv)
- MAXGEN generációs szám után az algoritmus leáll

- Induló populáció generálása véletlenszerűen (crtbp függvény)

### **SGA felépítése:**

```
NIND=40; %Number of individuals
MAXGEN=300; %Maximum no. of generations
NVAR=20; %No. of variables
PRECI=20; %Precision of variables
GGAP=0.9; %Generation map
```

#### %Build field descriptor

```
FieldD=[rep([PRECI],[1,NVAR]);rep([-512;512],[1,NVAR]);... rep([1;0;1;1],[1,NVAR])];
```

#### %Initialize population (crtbp:create binary population)

```
Chrom=crtbp(NIND,NVAR*PRECI);
gen=0; %Counter
```

#### %Evaluate initial population (bsrv: binary string to real value)

```
ObjV=objfun1(bs2rv(Chrom,FieldD));
```

#### %Generational loop

```
while gen<MAXGEN,
```

##### %Assign fitness values to entire population

```
FitnV=ranking(ObjV);
```

##### %Select individuals for breeding

```
SelCh=select('sus',Chrom,FitnV,GGAP);
```

##### %Recombine individuals (crossover)

```
SelCh=recombin('xovsp',SelCh,0.7);
```

##### %Apply mutation

```
SelCh=mut(SelCh);
```

##### %Evaluate offspring, call objective function

```
ObjVSel=objfun1(bs2rv(SelCh,FieldD));
```

##### %Reinsert offspring into population

```
[Chrom ObjV]= reins(Chrom,SelCh,1,1,ObjV,ObjVSel);
```

##### %Increment counter

```
gen=gen+1;
```

```
end
```

## 2. Átszámítás célfüggvényről fitness értékre, lineáris és nemlineáris rangsor. Szelekciós algoritmusok, rulett-kerék módszer, sztochasztikus univerzális mintavételezés (SUS).

**Kiindulás:** Az  $x = (x_1, \dots, x_{N_{var}}) \mapsto f(x)$  objektív függvény értéke, amely pozitív és negatív is lehet. Ezt kell minimalizálni!

**Elvárt:** Az  $F(x) = g(f(x))$  fitness függvény értéke legyen pozitív és őrizze meg a rendezettséget.

A rangsorolási függvény (ranking) az ObjV vektorban található  $N_{ind}$  darab egyed objektív függvény értékéhez egy fitness értéket rendel és eltárolja a FitnV vektorban.

Legyenek a populáció egyedei fenotípus (real) alakban az  $x_i$  vektorok:  $\{x_i\}_{i=1}^{N_{ind}}$ .

**Elvek:**

Pozitív objektív függvény érték esetén arányos fitness:

$$F(\mathbf{x}_i) = \frac{f(\mathbf{x}_i)}{\sum_{i=1}^{N_{ind}} f(\mathbf{x}_i)}$$

Skálázás (scaling) esetén:

- Kód:  $\text{FitnV} = \text{scaling}(\text{ObjV}, \text{Smult})$ ;
- Eljárás:  $F(x_i) = a f(x_i) + b$ 
  - o  $a$  és  $b$  értékét határozza meg a függvény úgy, hogy a maximális fitness felső korlátja az ObjV átlagértékének ( $f_{ave}$ ) Smult-szorosa legyen.

Rangsorolás (ranking) esetén:

- Kód:  $\text{FitnV} = \text{ranking}(\text{ObjV})$ ;
- Eljárás:
  - o  $f(x_i) \rightarrow pos$  hozzárendelés, ahol a legkisebb objektív függvény érték adja a max. fitnessst:

$$\max f(\mathbf{x}_i) \rightarrow pos(f(\mathbf{x}_i)) = 1, \dots, \min f(\mathbf{x}_i) \rightarrow pos(f(\mathbf{x}_i)) = N_{ind}$$

- o Ezután a rangsor lehet:
  - Lineáris

$$F(\mathbf{x}_i) = 2 - sp + 2 * (sp - 1) \frac{pos(f(\mathbf{x}_i)) - 1}{N_{ind} - 1}$$

pl.  $sp = 2$  estén a legjobb egyed 2, a legrosszabb 0 fitness értéket kap

- Nemlineáris

- 1. Meghatározni a következő egyenlet X gyökeit:

$$(sp - N_{ind})x^{N_{ind}-1} + sp * x^{N_{ind}-2} + \dots + sp * x + sp = 0$$

- Meghatározni az egyedek fitness értékét:

$$F(\mathbf{x}_i) = \frac{N_{ind} * X^{pos(f(\mathbf{x}_i)) - 1}}{\sum_{i=1}^{N_{ind}} X^{i-1}}$$

Látható, hogy a fitness értéket nem az objektív függvény értékhez, hanem annak pozíciójához rendeljük, így megakadályozható az idő előtti konvergencia egy lokális minimumhoz.

### Szelekciós algoritmusok:

A szelekció annak a kísérletszámnak a meghatározása, hogy egy egyed hányszor kerüljön kiválasztásra a rekombinációhoz (kereszteléshez). Eldönti, hogy egy egyed hány utód létrehozásában vegyen részt.

A szelekció két független folyamatból áll:

1. Kísérletszám meghatározása, amelyre egy egyed számíthat. (Fitness → reprodukcióban való részvétel valószínűsége)
2. A kísérletek számának konvertálása diszkrét számú utóddá. (Reprodukcióban való részvétel valószínűsége → Az egyedek tényleges szelekciója reprodukcióra)

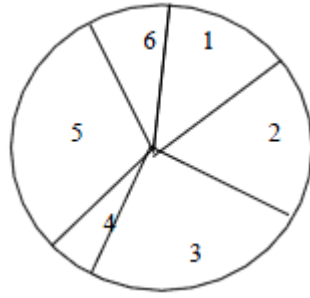
A szelekciós folyamat jellemzői:

- Torzítás (bias): Az egyed aktuális és várható szelekciós valószínűsége közötti különbség → pontosság
- Szórás (spread): A lehetséges kísérletek értéktartománya, amelyben egy egyed részt vehet → konzisztencia

Cél a hatékonysághoz: nulla torzítás, minimális szórás, kis időigény.

### Rulett kerék szelekció (roulette wheel selection):

1. Az egyedeket a  $[0, Sum]$  intervallumba képezzük le, ahol  $Sum$  az egyedek fitness értékeinek az összege.
2. Az egyedek intervallum mérete az egyedek fitness értékeinek felel meg.
3. Az egyed szelektálása egy egyenletes eloszlású véletlen számmal  $[0, Sum]$  között.
4. Annyi véletlen számot generálunk, amennyi egyedet kívánunk kisselektálni.



### Sztocasztikus univerzális mintavételezés (Stochastic universal sampling - SUS)

Ellentétben a rulett-kerék szelekciós módszertől nem egy, hanem  $N$  pontert alkalmazunk egyszerre, azaz az összes kiválasztandó egyedet egy fázisban szelektáljuk ki.

1. A populációt véletlenszerűen megkeverik (shuffle)
2. Egy véletlenszám generálása  $[0, Sum/N]$  között.
3. Az  $N$  egyed ezután szelektálva  $\left\{ptr, ptr + \frac{Sum}{N}, \dots, ptr + (n - 1) \frac{Sum}{N}\right\}$  pointerekkel.

Jellemzői: Minimális szórás, nulla torzítás  $\rightarrow$  a legjobb eredmény!

Kód: `SelCh=select('sus',Chrom,FitnV,GGAP);` ('sus' – stochastic universal sampling, 'rws' – roulette wheel selection)

### 3. Genetikus operátorok (keresztezés, mutáció) és megvalósításuk bináris és real GA esetén. Visszahelyettesítési stratégiák.

#### Rekombináció/Keresztezés

Az új kromoszómák (egyedek) generálásának módja a keresztezés (crossing).

A szelekció után következik. A kiszelektált egyedeket párokba soroljuk és ezek között lesz a keresztezés.

A rekombináció során keletkező egyedek populációba való visszahelyettesítése (reinsertation) történhet:

- Véletlenszerűen
- Fitness értékük figyelembevételével

#### Rekombináció bináris populáció esetén

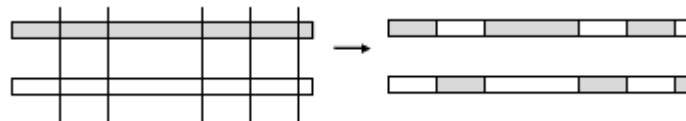
Legyen  $L$  a bináris kromoszóma string hossza.

#### Egypontos keresztezés

1. Egyenletes eloszlású véletlen  $i$  szám generálása  $[1, L-1]$  között minden páratlan/páros sorszámú kromoszómapárhoz.
2. Az  $i$  indexű bittől kezdve a biteket kicseréljük a két kromoszóma között.

### Multipontos keresztezés

1.  $M$  keresztezési pont kiválasztása véletlenszerűen:  $k_i \in \{1, 2, \dots, L-1\}$
2. Az első lépést minden kiszelektált páratlan/páros kromoszómapárhoz elvégezzük.
3. A keresztezési pontokat növekvő sorrendbe rendezzük a kromoszómapárnál, és a páratlan-páros kereszteződési pontok között a biteket kicseréljük. Így két új kromoszóma keletkezik.



### Egyenletes keresztezés

Minden kromoszóma hely keresztezési ponttá válhat.  $L$  hosszúságú maszkot generálunk minden kromoszómapárhoz, amely megmondja, mely biteket cseréljük ki.

### **Rekombináció real populáció esetén**

A bináris populációk keresztezési operátorai nem alkalmazhatók.

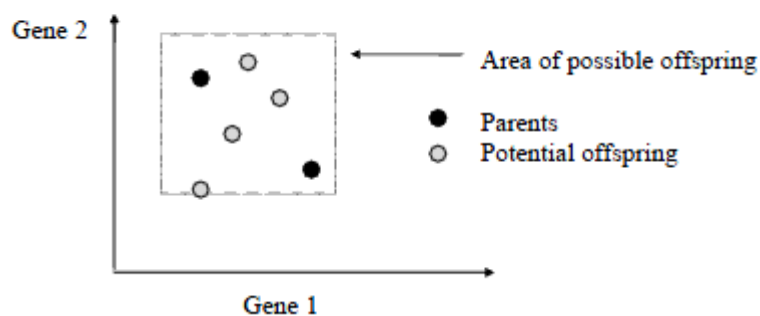
Az utódok fenotípusa real populáció esetén a páratlan/páros sorszámú szülők fenotípus értékei körül és között jelentkeznek.

Típusai:

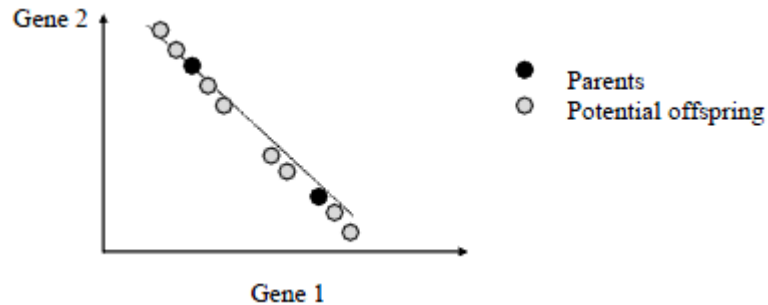
- Közbenső rekombináció (intermediate recombination)
  - o Az utódok  $O_1$  és  $O_2$  fenotípus értékei a szülők  $P_1$  és  $P_2$  fenotípus értékeiből egy skálatényező figyelembevételével keletkeznek, ahol  $\mathbf{a}$  egy véletlenvektor, aminek komponenseit pl. a  $[-0.25, 1.25]$  intervallumból választjuk.

$$\sigma = P_1 + \mathbf{a} * (P_2 - P_1)$$

$$\dim \mathbf{a} = N_{\text{var}}$$



- Egyenes mentén történő rekombináció (line recombination)
  - o  $\sigma = P_1 + \alpha \cdot (P_2 - P_1)$ , ahol  $\alpha$  egy véletlen szám. Mindez a közbelső rekombináció azon elfajuló esete, ahol az a vektor minden komponense azonos, függetlenül a fenotípus változó indexétől.



### Mutáció

Véletlen folyamat, amelynek keretében egy allélja egy másikkal helyettesítődik az új genetikai struktúra létrehozása érdekében.

Véletlenszerűen és kis valószínűséggel (0.001-0.01) alkalmazzuk.

Cél: Egyetlen string keresési valószínűsége se legyen nulla és ne vesszen el a jó genetikai anyag a szelekció és a keresztezés következtében.

Kód: `SelCh=mut(SelCh);`

### Bináris populáció esetén:

Mutation point ↓		binary	Gray
Original string-	0 0 0 1 1 0 0 0 1 0	0.6634	0.9659
Mutated string-	0 0 1 1 1 0 0 0 1 0	1.8439	2.2146

Kis változtatás, jelentős változás!

A Gray-kód előnyösen alkalmazható a genetikai algoritmusokban, mivel a szabályos Hamming távolság a kvantálási intervallumok között a sztochasztikus keresést kevésbé téveszti meg.

### Valós populáció esetén:

Megvalósítás:

- A gén értékek perturbációja vagy
- Az új értékek véletlenszerű megváltoztatása

$$x_{mut} = x + Mutmask * \frac{U - L}{2} * Mutshrink * Delta$$

- Mutshrink: Összehúzási paraméter
- Mutmask: Maszk, amely a változót egy *Mutate* (pl. ) mutációs ráta alapján jelöli ki mutációra. Értéke: -1,0 vagy 1
- Delta =  $\sum_{i=1}^{m-1} \alpha_i 2^{-i}$ , ahol  $m$  a pontosság (bitszám)

Ez a mutációs operator pl.  $m = 20$  esetén képes az optimum helyét  $0.5(U - L)Mutshrink \cdot 2^{-19}$  pontossággal megtalálni.

### Visszahelyettesítés

Miután a régi populációból létrejött az új generáció a szelekció és a rekombináció révén, meghatározható az új populáció egyedeinek értéke.

Ha kevesebb egyed keletkezett a rekombináció során, akkor a régi és az új populáció közötti különbség egy generációs űrt (*generation gap*) definiál.

- Ha csak egy vagy két új egyed keletkezik  $\rightarrow$  *Állandósult állapotú (vagy inkrementális) GA*.
- Ha a legalkalmasabb egyedek generációról generációra fennmaradhatnak  $\rightarrow$  *Elitista stratégia*.

Az utódok visszahelyettesítésre való kiválasztása történhet:

- Véletlenszerűen (uniform selection)
- Fitness értékek alapján (fitness based selection)

Az utódok visszahelyettesítési aránya előírható (a populáció, ill mpga esetén a szubpopuláció méretének arányában)

## 4. Multipopulációs genetikus algoritmus (MPGA) blokkvázlata, migrációs stratégiák.

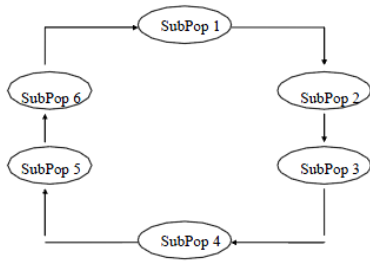
### Migráció

Multipopulációs esetben a populáció szubpopulációkból áll. Az új szubpopuláció mindig a megfelelő régi szubpopulációba kerül visszahelyettesítésre. A különböző szubpopulációk közötti egyedátadás a migráció.

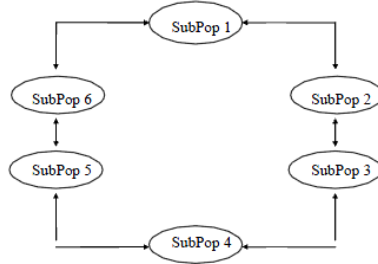
Megadása:

- Migrációs ráta (szubpopuláció méretének százalékában)
- Migrációra kiválasztás típusát (véletlenszerű vagy fitness alapú)
  - o Migrációs topológia:

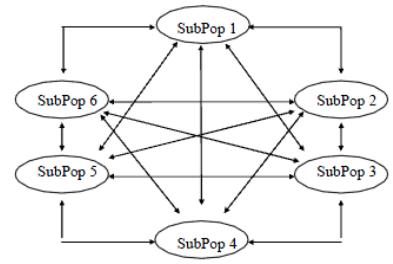




Gyűrű alakú topológia



Szomszédos topológia



Korlátatlan topológia