

Metody Rozmyte i Algorytmy Ewolucyjne

mgr inż. Piotr Kaczyński

Wydział Matematyczno-Przyrodniczy
Szkoła Nauk Ścisłych
Uniwersytet Kardynała Stefana Wyszyńskiego

Kodowanie
Operatory genetyczne

Plan wykładu

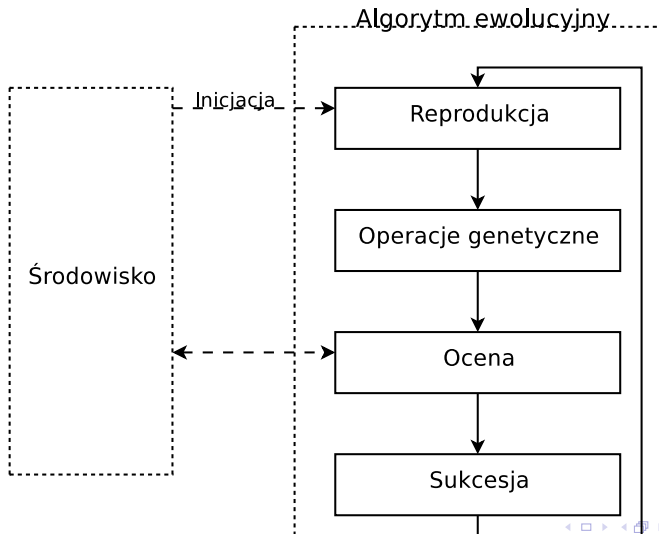
- 1 **Kodowanie**
 - Genotyp i fenotyp
 - Cechy kodowania
- 2 **Operatory genetyczne**
 - Pożądane cechy operatorów
 - Operatory krzyżowania
 - Krzyżowanie wymieniające
 - Operatory krzyżowania uśredniającego
 - Porównanie
 - Krzyżowanie wieloosobnicze
 - Operatory mutacji
- 3 **Eksploracja i eksploatacja a operatory**
 - Modyfikacja zasięgu mutacji
 - Modyfikacja zasięgu krzyżowania

Ogólny schemat algorytmu ewolucyjnego

- $t = 0$
- Inicjacja P^t
- Ocena P^t
- while not warunek stopu do
 - $T^t =$ reprodukcja P^t
 - $O^t =$ operacje genetyczne T^t
 - ocena O^t
 - $P^{t+1} =$ sukcesja(P^t, O^t)
 - $t = t + 1$



Schemat algorytmu ewolucyjnego - przypomnienie



Genotyp i fenotyp

- **Genotyp** osobnika, to jego chromosomy, zestaw liczb,
- **Fenotyp** osobnika, to rozwiązanie, które on przechowuje,
- Genotyp jest „przepisem” na utworzenie fenotypu,
- Mówimy o **kodowaniu** fenotypu przez genotyp,
- Genotyp jest punktem w **przestrzeni kodów**, fenotyp jest punktem w **przestrzeni rozwiązań** problemu,
- Często oba są tożsame,

Przykład

Minimalizacja funkcji jednej zmiennej $x \in \mathbb{Z}^n$. Przestrzeń fenotypów to \mathbb{Z}^n , przestrzeń genotypów można zdefiniować dwojako: jako \mathbb{Z}^n (kodowanie naturalne) lub np. jako przestrzeń $\{0, 1\}^{32}$ liczb w reprezentacji binarnej. W drugim przypadku chromosom składa się z 32 genów.



Znaczenie kodowania

- Wybór kodowania jest fundamentalnym do rozwiązania problemu,
- Wybierając sposób kodowania należy wziąć pod uwagę
 - Łatwość dekodowania,
 - Szybkość obliczenia funkcji celu na genotypie,
 - Weryfikacja spełnienia ograniczeń,

Przykład

Zastosowanie kodowania binarnego liczby całkowitej z poprzedniego przykładu pozwala w naturalny sposób uwzględnić ograniczenie

$$0 \leq x \leq 2^n$$

poprzez dobór odpowiedniej długości chromosomu.



Pożądane cechy kodowania

- Niech x będzie punktem w przestrzeni fenotypu,
- Niech X będzie punktem w przestrzeni genotypu,
- Poprzez F oznaczmy funkcję dekodującą, to jest

$$x = F(X)$$

- Kodowanie powinno być tak dobrane, aby

$$\forall x \exists X \quad F(X) = x, \quad (1)$$

$$|x^1 - x^2| \geq |x^1 - x^3| \implies |X^1 - X^2| \geq |X^1 - X^3| \quad (2)$$



Operatory genetyczne

- Operatory genetyczne określają sposób przekształcania **genotypów**
- Mutacja to odwzorowanie

$$M : \mathcal{D} \rightarrow \mathcal{D}$$

- Krzyżowanie to odwzorowanie

$$K : \mathcal{D}^k \rightarrow \mathcal{D}^l$$

W krzyżowaniu bierze udział $k \geq 2$ **rodziców**, tworząc l potomków

- Jeśli $k \geq 3$ to mówimy o krzyżowaniu **wieloosobniczym**,
- **Operatory genetyczne są kluczowym elementem algorytmu genetycznego**



Spójność przestrzeni genotypów

- Podstawowa kwestia: **czy można wygenerować dowolny chromosom z dowolnego innego posługując się jedynie operatorami genetycznymi?**
- Należy zadbać o to, aby

$$\forall X \in P^t \forall Y \in D \exists \{X^i\}$$
$$X^{i+1} = \text{Operacje genetyczne}(X^i)$$
$$X, Y \in \{X^i\}$$

- Własność tą nazywamy **spójnością przestrzeni genotypów**
- Gwarantuje ona, że algorytm który został źle zainicjowany nie jest z góry skazany na niepowodzenie,
- Każdy z punktów przestrzeni genotypu jest osiągalny z pewnym (niezerowym) prawdopodobieństwem,



Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji.

Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji. **Nie spełnia postulatu spójności**



Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji. **Nie spełnia postulatu spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z krzyżowaniem jednopunktowym i z operatorem mutacji zmieniającej bit na przeciwny.

Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji. **Nie spełnia postulatu spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z krzyżowaniem jednopunktowym i z operatorem mutacji zmieniającej bit na przeciwny. **Spełnia postulat spójności**

Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji. **Nie spełnia postulatu spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z krzyżowaniem jednopunktowym i z operatorem mutacji zmieniającej bit na przeciwny. **Spełnia postulat spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z operatorem mutacji zmieniającej bit na przeciwny, bez operatora krzyżowania.



Spójność przestrzeni genotypów - przykład

Przykład

Algorytm genetyczny z kodowaniem binarnym, krzyżowaniem jednopunktowym i bez operatora mutacji. **Nie spełnia postulatu spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z krzyżowaniem jednopunktowym i z operatorem mutacji zmieniającej bit na przeciwny. **Spełnia postulat spójności**

Przykład

Algorytm genetyczny z kodowaniem binarnym, z operatorem mutacji zmieniającej bit na przeciwny, bez operatora krzyżowania. **Również spełnia postulat spójności**



Obciążoność operatorów

- Spełnienie postulatu spójności nie gwarantuje poprawnego działania algorytmu,
- Niekiedy pewne kierunki poszukiwań są bardziej prawdopodobne, niezależnie od funkcji przystosowania,

Definition

Mówimy, że operator genetyczny jest **obciążony** jeśli przy braku nacisku selektywnego, wystąpienie niektórych genotypów w populacjach jest bardziej prawdopodobne niż innych.



Obciążoność operatorów

- Spełnienie postulatu spójności nie gwarantuje poprawnego działania algorytmu,
- Niekiedy pewne kierunki poszukiwań są bardziej prawdopodobne, niezależnie od funkcji przystosowania,

Definition

Mówimy, że operator genetyczny jest **obciążony** jeśli przy braku nacisku selektywnego, wystąpienie niektórych genotypów w populacjach jest bardziej prawdopodobne niż innych.

- Dla mutacji: **duża zmiana genotypu mniej prawdopodobna niż mała,**
- Dla krzyżowania: **dalsze krzyżowanie chromosomów osiągalnych w wyniku krzyżowania nie powoduje powstania chromosomów nie należących do zbioru**



Warianty operatorów krzyżowania

- Para rodziców, para potomków
 - Najczęstszy wariant,
 - Chromosomy potomne Y , Z są **sprzężone**, to jest

$$Y = c(X^1, X^2),$$

$$Z = c(X^2, X^1)$$

- Pojedynczy osobnik potomny (krzyżowanie dwuosobnicze),
- Krzyżowanie globalne
 - Występuje jeden osobnik wiodący i n osobników pomocniczych,
 - Każdy osobnik pomocniczy odpowiada za jeden gen,



Krzyżowanie a licznosc populacji

- Do krzyżowania osobniki wybierane są przez reprodukcję,
- Licznosc populacji tymczasowej T^t jest zależna od liczby potomków,
- Bardzo często krzyżowanie występuje z pewnym **prawdopodobieństwem krzyżowania** p_c , który jest jednym z ważniejszych parametrów algorytmu,
- W praktyce operacja krzyżowania i reprodukcji mogą być wykonywane jednocześnie
 - Z populacji bazowej reprodukowany jest osobnik,
 - Jeśli zaszło zdarzenie, że zostanie poddany krzyżowaniu to wykonywana jest operacja reprodukcji tyle razy, aby dobrać odpowiednią ilość rodziców i wykonywane jest krzyżowanie
 - Operacje te powtarzane są tak długo, aż populacja tymczasowa osiągnie odpowiednią licznosc,



Krzyżowanie wymieniające

- Chromosomy potomne powstają przez składanie z chromosomów rodzicielskich,
- Jeśli liczba potomków jest równa liczbie rodziców to schemat ten nie doprowadzi do zmian zróżnicowania genów w populacji,
- Można je wykorzystywać zarówno w kodowaniu binarnym jak i rzeczywistoliczbowym,

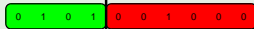
Przykład

Osobniki rodzicielskie

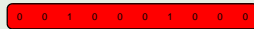
chromosom 1



chromosom 2



Osobniki potomne



Krzyżowanie jednopunktowe

- Wykonywane dwuetapowo
 - Najpierw wybierana jest liczba c ze zbioru $\{1, 2, \dots, n - 1\}$,
 - Liczba ta stanowi punkt rozcięcia,
 - Osobniki podzielone na dwie części są sklejane przy czym osobniki potomne mają po jednej części z każdego osobnika rodzicielskiego
- W schemacie z jednym osobnikiem potomnym

$$Y = [X_1^1, \dots, X_c^1, X_{c+1}^2, \dots, X_n^2]$$

- W schemacie z dwoma osobnikami potomnymi

$$Z = [X_1^2, \dots, X_c^2, X_{c+1}^1, \dots, X_n^1]$$

- Wymianie ulegnie wartość co najmniej jednego genu,



Krzyżowanie jednopunktowe

- Wykonywane dwuetapowo
 - Najpierw wybierana jest liczba c ze zbioru $\{1, 2, \dots, n - 1\}$,
 - Liczba ta stanowi punkt rozcięcia,
 - Osobniki podzielone na dwie części są sklejane przy czym osobniki potomne mają po jednej części z każdego osobnika rodzicielskiego
- W schemacie z jednym osobnikiem potomnym

$$Y = [X_1^1, \dots, X_c^1, X_{c+1}^2, \dots, X_n^2]$$

- W schemacie z dwoma osobnikami potomnymi

$$Z = [X_1^2, \dots, X_c^2, X_{c+1}^1, \dots, X_n^1]$$

- Wymianie ulegnie wartość co najmniej jednego genu,
- **Schemat ten jest obciążony.**



Krzyżowanie jednopunktowe - przykład

Przykład

Rozważmy dwa chromosomy rodzicielskie

$$X^1 = [000], \quad X^2 = [111]$$

Założmy powstawanie dwóch osobników potomnych. Mamy



Krzyżowanie jednopunktowe - przykład

Przykład

Rozważmy dwa chromosomy rodzicielskie

$$X^1 = [000], \quad X^2 = [111]$$

Założmy powstawanie dwóch osobników potomnych. Mamy

$$\text{Dla } c = 1 \quad Y = [011], \quad Z = [100]$$

$$\text{Dla } c = 2 \quad Y = [001], \quad Z = [110]$$

Zbiór osiągalny to $\{[001], [011], [100], [110]\}$, natomiast jeśli weźmiemy do krzyżowania pierwszą z par potomków to dla $c = 2$ otrzymamy $Y = [010], Z = [101]$ które nie należały do zbioru osiągalnego w pierwszym krzyżowaniu.



Krzyżowanie dwupunktowe

- Modyfikacja krzyżowania jednopunktowego,
- Losowane są dwa punkty rozcięcia c_1, c_2 będące dwuelementową kombinacją ze zbioru $\{1, 2, \dots, n - 1\}$
- Krzyżowane chromosomy dzieli się na **trzy** części, wymianie podlega środkowy
- W przypadku jednego osobnika potomnego

$$Y = [X_1^1, \dots, X_{c_1}^1, X_{c_1+1}^2, \dots, X_{c_2}^2, X_{c_2+1}^1, \dots, X_n^1]$$

- W przypadku drugiego osobnika potomnego

$$Y = [X_1^2, \dots, X_{c_1}^2, X_{c_1+1}^1, \dots, X_{c_2}^1, X_{c_2+1}^2, \dots, X_n^2]$$



Krzyżowanie dwupunktowe

- Modyfikacja krzyżowania jednopunktowego,
- Losowane są dwa punkty rozcięcia c_1, c_2 będące dwuelementową kombinacją ze zbioru $\{1, 2, \dots, n - 1\}$
- Krzyżowane chromosomy dzieli się na **trzy** części, wymianie podlega środkowy
- W przypadku jednego osobnika potomnego

$$Y = [X_1^1, \dots, X_{c_1}^1, X_{c_1+1}^2, \dots, X_{c_2}^2, X_{c_2+1}^1, \dots, X_n^1]$$

- W przypadku drugiego osobnika potomnego

$$Y = [X_1^2, \dots, X_{c_1}^2, X_{c_1+1}^1, \dots, X_{c_2}^1, X_{c_2+1}^2, \dots, X_n^2]$$

- Nadal jest to schemat obciążony!

Krzyżowanie dwupunktowe - wariant zmodyfikowany

- Modyfikacja sposobu losowania punktów c_1 oraz c_2 ,
- Losowanie ze zwracaniem,
- Możliwa jest sytuacja w której $c_1 = c_2$
- W takim wypadku stosowany jest schemat jednopunktowy,
- Dodatkowo prawdopodobieństwa osiągalności chromosomów potomnych są **różne**,

Przykład

Ponownie rozważmy $X^1 = [000]$, $X^2 = [111]$. Ponieważ miejsca cięcia są losowane ze zwracaniem, osobnik potomny $[010]$ może być utworzony, jeśli wylosowane zostaną $c_1 = 1$, $c_2 = 1$ lub $c_1 = 2$, $c_2 = 1$. Prawdopodobieństwo tego, że będzie on wynikiem krzyżowania jest więc większe.



Krzyżowanie dwupunktowe - wariant zmodyfikowany

- Modyfikacja sposobu losowania punktów c_1 oraz c_2 ,
- Losowanie ze zwracaniem,
- Możliwa jest sytuacja w której $c_1 = c_2$
- W takim wypadku stosowany jest schemat jednopunktowy,
- Dodatkowo prawdopodobieństwa osiągalności chromosomów potomnych są **różne**,

Przykład

Ponownie rozważmy $X^1 = [000]$, $X^2 = [111]$. Ponieważ miejsca cięcia są losowane ze zwracaniem, osobnik potomny $[010]$ może być utworzony, jeśli wylosowane zostaną $c_1 = 1$, $c_2 = 1$ lub $c_1 = 2$, $c_2 = 1$. Prawdopodobieństwo tego, że będzie on wynikiem krzyżowania jest więc większe.

- Jest to również schemat obciążony!



Krzyżowanie równomierne

- Chromosom potomny tworzony jest w następujący sposób

$$Y_i = \begin{cases} X_i^1 & \text{jeśli } \xi_{U(0,1),i} < p_e \\ X_i^2 & \text{w przeciwnym przypadku} \end{cases}$$

- Wartość p_e jest parametrem operacji krzyżowania,
- Typowo $p_e = \frac{1}{2}$,
- Jest to operator **nieobciążony**,
- W przypadku dwóch osobników potomnych

$$Z_i = \begin{cases} X_i^2 & \text{jeśli } Y_i = X_i^1 \\ X_i^1 & \text{w przeciwnym przypadku} \end{cases}$$



Operatory krzyżowania uśredniającego

- Oddziałują na wartości genów chromosomów poddawanych krzyżowaniu,
- Specyficzne dla kodowania rzeczywistoliczbowego,
- Wartości genów potomnych są z odcinka łączącego geny rodzicielskie,
- Dla schematów z parą osobników potomnych — chromosomy potomne są symetryczne względem środka odcinka łączącego osobniki rodzicielskie,
- Schematy te są nieobciążone



Wariant podstawowy

- Osobnik potomny generowany jest według wzoru

$$Y = X^1 + \xi_{U(0,1)}(X^2 - X^1)$$

- W przypadku dwóch osobników potomnych

$$Z = X^2 + X^1 - Y$$

- Każdy z osobników dopuszczalnych jest jednakowo prawdopodobny,
- Jeśli zbiór rozwiązań dopuszczalnych jest wypukły, to ten operator zawsze wygeneruje osobniki dopuszczalne



Wariant alternatywny

- Osobnik potomny generowany jest według wzoru

$$Y = X^1 + \xi_{U(0,1),i}(X^2 - X^1)$$

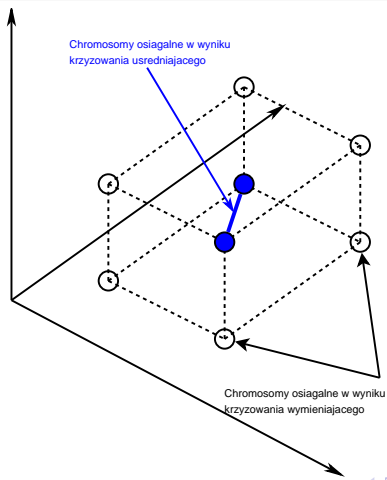
- W przypadku dwóch osobników potomnych

$$Z = X^2 + X^1 - Y$$

- Każdy z osobników dopuszczalnych jest jednakowo prawdopodobny,
- Zbiór osobników dopuszczalnych to wnętrze hiperprostopadłościanu, którego przeciwległymi narożnikami są osobniki rodzicielskie,



Krzyżowanie wymieniające i uśredniające - porównanie



Krzyżowanie wielosobnicze

- Zwiększenie ilości rodziców powoduje zwiększenie obszaru chromosomów osiągalnych,
- Krzyżowanie wymieniające
 - Chromosom potomny komponowany jest ze wszystkich genów chromosomów rodzicielskich,

$$Y_i = X_j^i, \text{ jeśli } j - 1 < h\xi_{U(0,1),i} \leq j$$

- Krzyżowanie uśredniające

$$Y = \sum_{j=1}^h a_j X^j, \quad \sum_{j=1}^h a_j = 1$$

gdzie wektor $[a_1, a_2, \dots, a_h]$ jest dobierany z rozkładem jednostajnym,



Mutacja w kodowaniu binarnym

- Dla każdego genu podejmuje się decyzję o jego modyfikacji z prawdopodobieństwem p_m
- Rozważa się dwie wersje

- Negacja bitu

$$Y_i = 1 - Y_i$$

- Zmiana bitu na losowy

$$Y_i = \xi_{B(0,1)}, i$$

- Schemat drugi jest równoważny pierwszemu z $p'_m = \frac{1}{2}p_m$



Mutacje przeznaczone do optymalizacji w \mathbb{R}^n

- Perturbacja chromosomu poprzez dodanie realizacji n -wymiarowej zmiennej losowej ξ_r

$$Y = X + \xi_r$$

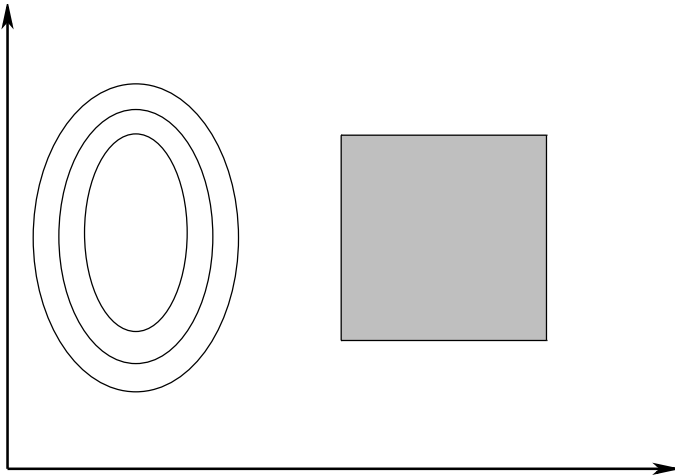
- Można również korzystać z rozkładów nieskorelowanych zmiennych jednowymiarowych

$$Y = X + \xi_{r,i}$$

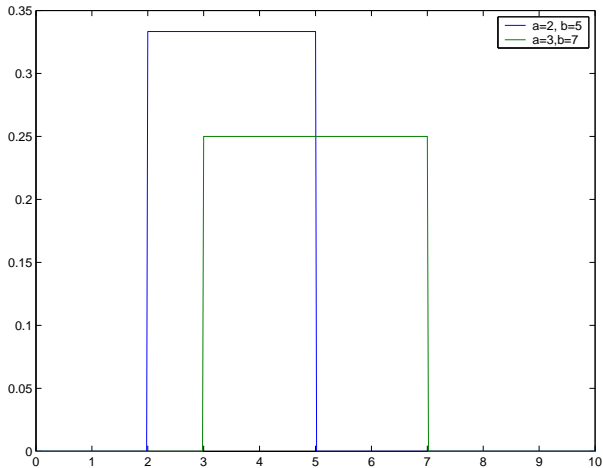
- Najczęściej wykorzystywane
 - Rozkład jednostajny,
 - Rozkład normalny,
 - Rozkład Cauchy'ego
- Każdy rozkład, dla którego $E[\xi_r] = 0$ jest nieobciążony



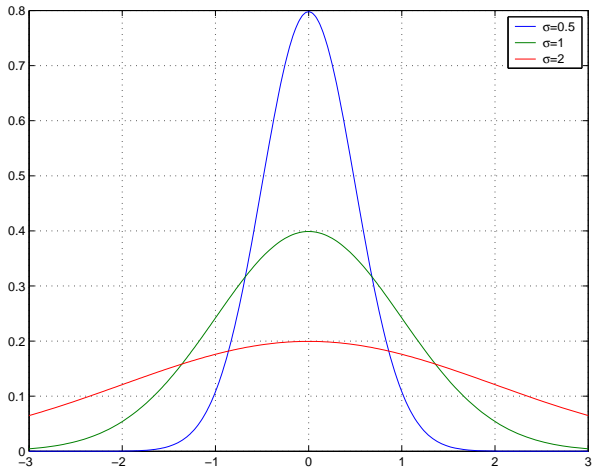
Obszary osiągalne



Rozkład jednostajny

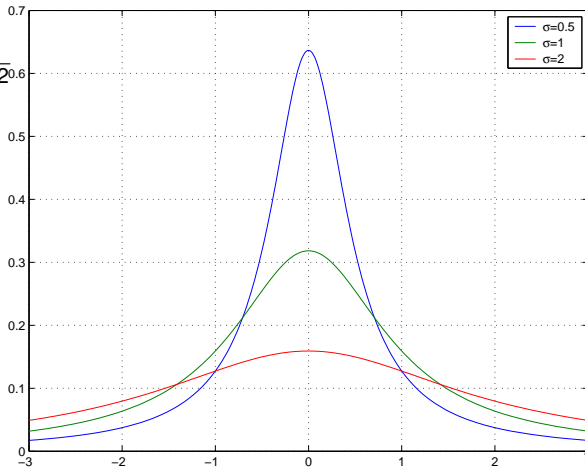


Rozkład normalny



Rozkład Cauchy'ego

$$f(x) = \frac{1}{\pi} \frac{\sigma}{\sigma^2 + (x - m)^2}$$



Rozkład normalny a rozkład Cauchy'ego

- W rozkładzie normalnym każda z wartości ma niezerową wartość gęstości prawdopodobieństwa,
- W praktyce wylosowanie liczby spoza przedziału $[-5\sigma, 5\sigma]$ jest niemożliwe,
- Wynika to z błędów reprezentacji numerycznej na maszynach cyfrowych,
- Rozkład Cauchy'ego ma **grubsze ogony**,
- Wylosowanie wartości odległych od środka symetrii jest większe niż dla rozkładu normalnego,
- **Rozkład Cauchy'ego nie posiada wartości oczekiwanej ani wariancji**



Eksploracja i eksploatacja a operatory genetyczne

- Który operator jest ważniejszy? **Nie wiadomo**,
- Operator mutacji gwarantuje spójność, operator krzyżowania rzadko,
- Zasięg operatora to pewna umowna miara zbioru chromosomów osiągalnych w wyniku działania tego operatora,
- Duży zasięg \implies silna eksploracja,
- Mały zasięg \implies silna eksploatacja,



Zasięg operatorów

- Zasięg mutacji
 - O zasięgu decyduje rozkład zmiennej zaburzającej,
 - Najczęściej zależy od wariancji tej zmiennej,
 - Duży zasięg mutacji to duże prawdopodobieństwo opuszczenia obszaru przyciągania optimum lokalnego,
- Zasięg krzyżowania
 - Sterowanie zasięgiem zależy od sposobu wyboru osobników rodzicielskich,
 - Mały zasięg otrzymujemy, gdy wybierane są osobniki należące do tego samego obszaru przyciągania maksimum lokalnego,
 - Duży zasięg jeśli osobniki te należą do różnych obszarów,
- Zasięg operatorów jest **względny** w stosunku do funkcji przystosowania i zależy od wielkości obszarów przyciągania poszczególnych minimów lokalnych,



Strojenie według scenariusza

- Zakłada się pewien scenariusz zmian zasięgu mutacji,
- Zasięg ten powinien być coraz mniejszy,
- Informacja o stanie algorytmu nie jest brana pod uwagę,

Przykład

Algorytm wykorzystujący kodowanie binarne i mutację zmieniającą bit na przeciwny z prawdopodobieństwem p_m

$$p_m(t) = ap_m(t - 1)$$

gdzie $0 \leq a \leq 1$ jest parametrem adaptacji



Adaptacja zasięgu

- Zmiana zasięgu zależy od pewnego wskaźnika opisującego algorytm,
- Wskaźnikiem tym może być np. różnorodność populacji,
- Przykłady
 - Reguła 1/5 sukcesów,
 - Zmniejszenie zasięgu wraz ze zmniejszaniem różnicy przystosowania osobnika i oszacowaniem wartości maksimum przystosowania



Samoczynna adaptacja

- Adaptacja realizowana przez dodatkowy chromosom,
- Dodatkowy chromosom nie ma wpływu na ocenę osobnika,
- Ulega on natomiast operacjom genetycznym,
- Tym samym zasięg operatorów jest dobierany przez ewolucję — promowane będą te osobniki, których zasięg operatorów jest lepszy,
- W przypadku kodowania binarnego zmieniane jest prawdopodobieństwo p_m ,



Modyfikacja zasięgu krzyżowania

- Rzadziej stosowana w praktyce,
- Zasięg krzyżowania można **zmniejszać** wraz z upływem czasu trwania algorytmu,
- Częściej zasięg krzyżowania jest jednak **powiększany**,
- Zmianę zasięgu można osiągnąć np. przez
 - Promowanie (lub zabranianie) krzyżowania osobników podobnych,
 - Modyfikację prawdopodobieństw osiągalności



Modyfikacje ze scenariuszem

- Zakaz krzyżowania osobników, których odległość genotypów przekracza próg c ,
- Wartość tą można modyfikować w czasie,
- Zakaz krzyżowania osobników, które miały wspólnego przodka w ostatnich k generacjach,
 - Konieczne trzymanie drzewa genealogicznego wszystkich osobników,
 - Jeśli występuje intensywna mutacja, to jest to dość skuteczna metoda,



Różnicowanie prawdopodobieństw osiągalności

Przykład

Krzyżowanie uśredniające, osobnik potomny jest generowany nie z losowaniem z przedziału $(0, 1)$ a z przedziału $(0, c)$.

$$Y = X^1 + c\xi_{U(0,1)}(X^2 - X^1)$$

gdzie c jest stosunkiem wartości funkcji gęstości rozkładu normalnego wykorzystywanego do mutacji chromosomu X^1 w chromosomie X^2 do wartości tego rozkładu w chromosomie X^1

$$c = \frac{f_{N(X^1, \sigma^1)}(X^2)}{f_{N(X^1, \sigma^1)}(X^1)}$$

