Metody Rozmyte i Algorytmy Ewolucyjne

mgr inż. Piotr Kaczyński

Wydział Matematyczno-Przyrodniczy
Szkoła Nauk Ścisłych
Uniwersytet Kardynała Stefana Wyszyńskiego

Strategie ewolucyjne
Podstawowe operatory genetyczne

Plan wykładu

1. Przypomnienie
   - Metody generacji liczb losowych
   - Schemat algorytmu ewolucyjnego
2. Ciągłe operatory genetyczne
   - Mutacja
   - Krzyżowanie
   - Inicjacja
3. Strategie ewolucyjne
   - Strategia $(1 + 1)$
   - Strategia $(\mu + \lambda)$
   - Strategia (μ, λ)

Generacja liczb losowych w komputerze

- Liczby są tylko pseudolosowe,
- Dostępny jest tylko generator liczb z rozkładu jednostajnego,
- Korzystając z tego generatora można wygenerować pozostałe rozkłady,
- Dla rozkładów dyskretnych podział odcinka $(0, 1)$,
- Dla rozkładów ciągłych wykorzystujemy metodę odwrotności dystrybuanty

Dystrybuanta i gęstość zmiennej losowej

Definition

Dystrybuantą rozkładu zmiennej losowej $X = [X_1, \ldots, X_n]^T$ nazywamy funkcję $F \colon \mathbb{R}^n \to \mathbb{R}$ daną wzorem

$$F(x_1, \ldots, x_n) = P(X_1 < x_1, \ldots, X_n < x_n)$$

Definition

Gęstością rozkładu zmiennej losowej $X = [X_1, \ldots, X_n]^T$ nazywamy funkcję $f \colon \mathbb{R}^n \to \mathbb{R}$ daną wzorem

$$f(x_1, \ldots, x_n) = F'(x_1, \ldots, x_n)$$

Rozkłady dyskretne

- Przyjmują wartości ze skończonego zbioru liczb $x_1, \ldots, x_n$,
- Każda wartość ma przyporządkowane prawdopodobieństwo $p_1, \ldots, p_n$,
- Oczywiście $\sum_{i=1}^{n} p_i = 1$,
- Metoda generacji liczb z rozkładu dyskretnego
  - Każdej wartości x_k przyporządkuj liczbę

$$F(x_k) = \sum_{i=1}^{k} p_i$$

  - Wylosuj liczbę u z rozkładu jednostajnego,
  - Znajdź najmniejsze takie k, dla którego $u \leqslant F(x_k)$,
  - Zwróć znalezione x_k

Generacja liczb z rozkładu dyskretnego

Przykład

Niech dany będą liczby $x_1 = 2$, której prawdopodobieństwo wypadnięcia jest $p_1 = \frac{1}{3}$, $x_2 = 7$ dla której prawdopodobieństwo wynosi $p_2 = \frac{1}{2}$ oraz $x_3 = 9$, dla której $p_3 = \frac{1}{6}$. Mamy

$$F(x_1) = \frac{1}{3}$$
$$F(x_2) = \frac{5}{6}$$
$$F(x_3) = 1$$

Jeśli z rozkładu jednostajnego na $(0, 1)$ wylosujemy np. $\frac{1}{4}$ to z modelowanego rozkładu zwrócimy $x_1 = 2$. Jeśli rozkład jednostajny zwróci na przykład $\frac{3}{4}$ to my zwrócimy $x_2 = 7$ itd.

Metoda odwrotności dystrybuanty

Theorem

Niech U będzie zmienną losową o rozkładzie jednostajnym na odcinku $[0, 1]$. Niech ponadto $F(x)$ będzie dystrybuantą pewnej zmiennej losowej X dla której istnieje funkcja odwrotna $F^{-1}(y)$. Wtedy zmienna

$$Y = F^{-1}(U)$$

ma rozkład identyczny z rozkładem zmiennej X.

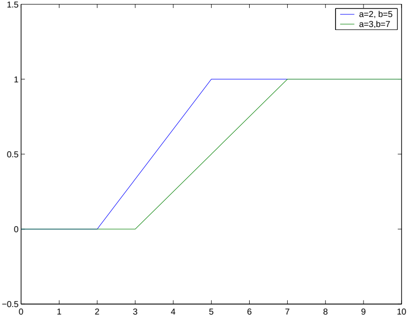
- Jeśli można policzyć funkcję odwrotną do zadanej dystrybuanty, to można generować liczby z tego rozkładu,
- Nie we wszystkich przypadkach da się znaleźć taką funkcję,

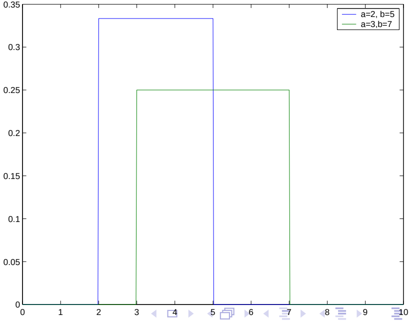
Rozkład jednostajny

- Gęstość i dystrybuanta rozkładu na odcinku $[a, b]$,

$$f(x) = \frac{1}{b-a} \quad \text{dla } x \in [a, b]$$

$$F(x) = \frac{1}{b-a}x \quad \text{dla } x \in [a, b]$$





Rozkład jednostajny — metoda generacji

- Jeśli U ma rozkład jednostajny na odcinku $[0, 1]$, to

$$Y = (b - a)U + a$$

ma rozkład jednostajny na odcinku $[a, b]$

Przykład

Generacja liczb losowych z odcinka $[7, 10]$

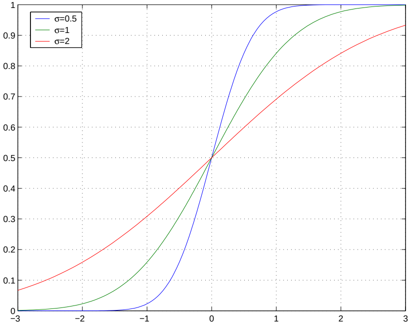
```
u = rand();
y = (10-7)*u + 7;
return y;
```

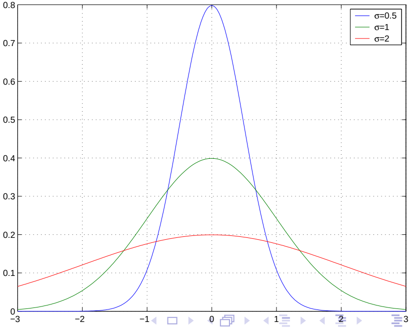
Rozkład normalny

- Gęstość dla $E[X] = \mu$, $Var[X] = \sigma^2$

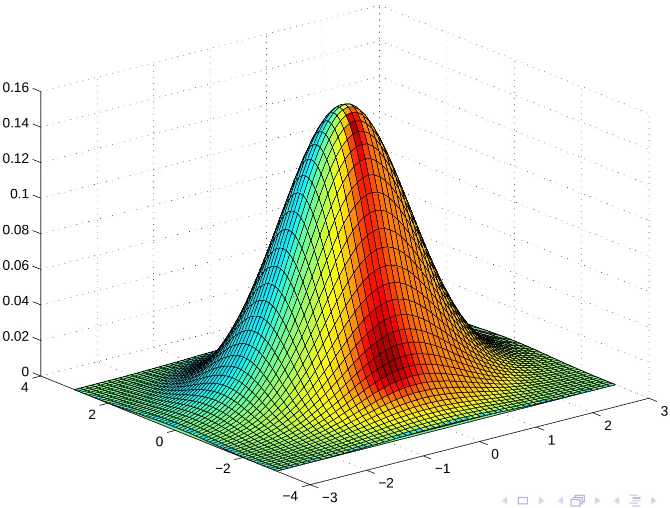
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

- Dystrybuanta nie da się wyrazić analitycznie





Dwuwymiarowy rozkład normalny



Generacja liczb ze std. rozkładu normalnego

- Standardowy rozkład normalny: $\sigma = 1$, $\mu = 0$,
- Transformata Boxa-Mullera,

Theorem

Jeśli U_1 oraz U_2 są niezależnymi zmiennymi losowymi o rozkładzie jednostajnym na $[0, 1)$, to

$$Y_1 = \sqrt{-2 ln U_1} \cos(2\pi U_2)$$
$$Y_2 = \sqrt{-2 ln U_1} \sin(2\pi U_2)$$

są niezależne i mają standardowy rozkład normalny

- Aby wygenerować wielowymiarowy, nieskorelowany rozkład normalny wystarczy kilkukrotnie wygenerować wartość z jednowymiarowego rozkładu normalnego/

Generacja liczb z rozkładu normalnego

- Korzystamy z własności wartości oczekiwanej i wariancji,

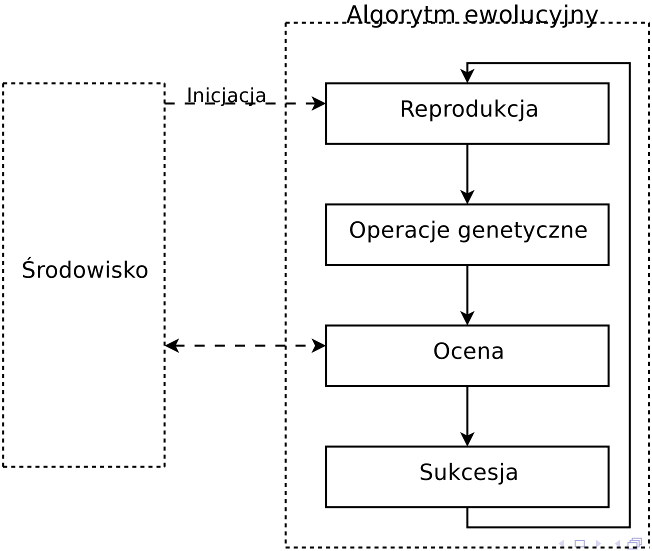
$$E[aX] = aE[X]$$
$$Var[aX] = a^2 Var[X]$$

- Aby wygenerować liczbę z rozkładu normalnego o μ oraz σ wystarczy wzór

$$Y = \sigma\xi + \mu$$

gdzie *xi* ma standardowy rozkład normalny

Schemat algorytmu ewolucyjnego



Etapy algorytmu

- Inicjacja — proces tworzenia nowych osobników, inicjowania ich chromosomów i obliczania funkcji przystosowania,
- Reprodukcja — kopiowanie osobników z uwzględnieniem wartości funkcji przystosowania,
- Operacje genetyczne — mutacja (jeden osobnik) i krzyżowanie (wiele osobników),
- Ocena — obliczenie funkcji przystosowania dla nowych osobników,
- Sukcesja — wybór odpowiednich osobników do następnej epoki,

Zadanie rzeczywistoliczbowe

- Poprzednio rozważano chromosomy z genami binarnymi,
- Krzyżowanie wymieniające,
- Mutacja zmieniająca bit na przeciwny,
- Co zrobić w przypadku chromosomów $x \in \mathbb{R}^n$?
- Wprowadzamy inne operatory krzyżowania i mutacji

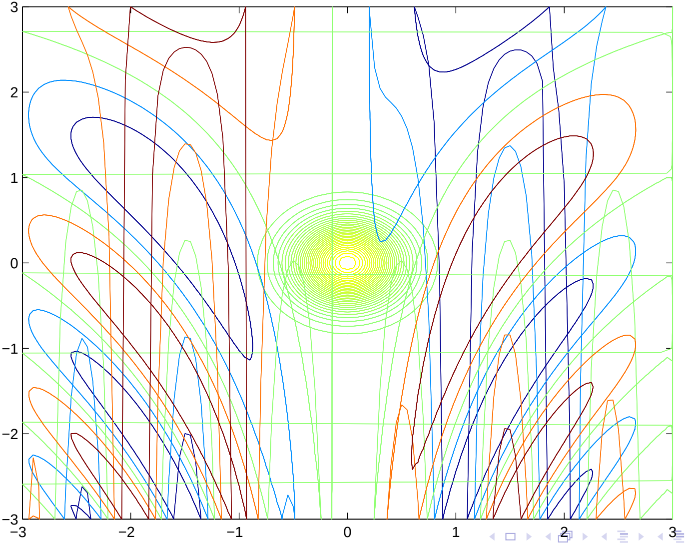
Mutacja rozkładem normalnym

- Mutacja jest losowym zaburzeniem genów,
- Losowa modyfikacja rozkładem normalnym

$$Y_i = X_i + \sigma\xi_i$$

  gdzie Y_i to i-ty gen po mutacji, X_i jest i-tym genem przed mutacją, ξ to zmienna losowa o standardowym rozkładzie normalnym,
- Parametr σ modyfikuje zasięg mutacji,
- Małe zmiany są bardziej prawdopodobne niż duże,
- Losowanie odbywa się dla każdego genu osobno.

Mutacja rozkładem normalnym II



Krzyżowanie uśredniające

- Krzyżowanie dwuosobnikowe,
- Powstaje jeden osobnik, który jest średnią z obu rodziców

$$Y = \lambda X^1 + (1 - \lambda) X^2$$

gdzie Y osobnik potomny, X^1 oraz X^2 osobniki rodzicielskie, natomiast λ jest zmienną losową o rozkładzie jednostajnym na $(0, 1)$,
- Osobnik potomny leży w losowym punkcie odcinka łączącego osobniki rodzicielskie,
- Zakładamy intuicyjnie, że pomiędzy dwoma dobrymi rozwiązaniami leży trzecie, jeszcze lepsze,
- Bardzo dobry do funkcji wypukłych,

Inicjacja w $\mathbb{R}^n$

- Problem: przestrzeń $\mathbb{R}^n$ jest nieskończonej miary,
- W praktyce: ograniczamy przestrzeń w której mogą znaleźć się początkowo osobniki,
- Najczęściej ograniczenia kostkowe

Przykład

Jeśli założymy, że początkowo każdy gen ma spełniać

$$a \leqslant x_i \leqslant b$$

to każdy gen w każdym chromosomie należy wybrać z rozkładem jednostajnym na odcinku $[a, b]$,

Wstęp

- Jedno z najstarszych podejść do algorytmów genetycznych,
- Stosują typowe operatory krzyżowania i mutacji,
- Są działającymi i sprawdzonymi algorytmami genetycznymi,
- Nie trzeba dobierać każdego operatora z osobna

Strategia $(1+1)$

- Pierwsza strategia ewolucyjna,
- Przetwarza tylko jednego osobnika z jednym chromosomem,
- Brak krzyżowania,
- Mutacja z zaburzeniem rozkładem normalnym

$$Y_i = X_i + \sigma\xi_i$$

- Zasięg mutacji zmienny w czasie — reguła $1/5$ sukcesów

Schemat strategii $(1+1)$

- $t = 0$
- inicjacja X^t
- Ocena X^t
- `while not` **warunek stopu** `do`
  - $Y^t =$ mutacja X^t
  - ocena Y^t
  - `if` $f(Y^t) > f(X^t)$ `then`
    - $X^{t+1} = Y^t$
  - `else`
    - $X^{t+1} = X^t$
- $t = t + 1$

Reguła $1/5$ sukcesów

- Reguła stosowana do zmiany parametru σ,
- Jeśli przez kolejnych k generacji liczba mutacji zakończonych sukcesem jest większa niż $1/5$ ogólnej liczby wykonanych mutacji, to należy zwiększyć zasięg mutacji
$$\sigma' = c_i\sigma, \quad c_i > 1$$
- Gdy dokładnie $1/5$ mutacji kończy się sukcesem, wartość σ pozostaje ta sama,
- W przeciwnym wypadku należy zawęzić zasięg mutacji
$$\sigma' = c_p\sigma, \quad c_p < 1$$
- Adaptacja mutacji pozwala na dokładniejsze przeszukiwanie optimów lokalnych

Strategia $(\mu+\lambda)$

- Jedna z najczęściej wykorzystywanych w praktyce,
- Uogólnienie strategii $(1+1)$, rozważana jest populacja osobników,
- Samoczynny mechanizm adaptacji zastępujący regułę $1/5$ sukcesów,
- Wprowadzono operator krzyżowania,
- Kodowanie osobników dwuchromosomowe
  - Chromosom X z rozwiązaniem zadania,
  - Chromosom σ z wariancjami dla mutacji poszczególnych genów,

Schemat strategii $(\mu + \lambda)$

- $t = 0$
- Inicjacja P^t
- Ocena P^t
- `while not` warunek stopu `do`
    - T^t = reprodukcja P^t
    - O^t = krzyżowanie i mutacja T^t (λ osobników)
    - ocena O^t
    - $P^{t+1} = \mu$ najlepszych osobników z $P^t \cup O^t$
    - $t = t + 1$

Mutacja

- Losowana jest wartość ψ ze standardowego rozkładu normalnego,
- Każdy element chromosomu wariancji σ jest modyfikowany wg wzoru
$$\sigma_i' = \sigma_i \exp\left(\tau' \psi + \tau \xi_i\right)$$
gdzie ξ jest zmienną losową o standardowym rozkładzie normalnym, a τ i τ' są parametrami,
- Mutowany jest chromosom X przy użyciu chromosomu wariancji σ
$$X_i' = X_i + \sigma_i' \xi_i$$
gdzie $\xi_i \sim N(0, 1)$
- Typowe wartości parametrów τ oraz τ' to
$$\tau = \frac{K}{\sqrt{2n}}, \quad \tau' = \frac{K}{\sqrt{2\sqrt{n}}}$$

Krzyżowanie i reprodukcja

- Krzyżowanie uśredniające z zastosowaniem tego samego współczynnika dla wszystkich chromosomów
- Dwa osobniki potomne i dwa rodzicielskie

$$
\begin{aligned}
X'^1 &= \rho X^1 + (1-\rho)X^2 \\
X'^2 &= \rho X^2 + (1-\rho)X^1 \\
\sigma'^1 &= \rho\sigma^1 + (1-\rho)\sigma^2 \\
\sigma'^2 &= \rho\sigma^2 + (1-\rho)\sigma^1 \\
\rho &\sim U(0,1)
\end{aligned}
$$

- Reprodukcja — proste losowanie λ osobników ze zwracaniem wśród μ osobników populacji bazowej

Strategia (μ, λ)

- Modyfikacja strategii $(\mu+\lambda)$,
- Problemy z bardzo silnymi osobnikami,
- Osobniki takie „nie umierają”,
- Wpływa to na stagnację populacji
- Każdy osobnik w strategii (μ, λ) żyje dokładnie jedną epokę

Schemat strategii (μ, λ)

- $t = 0$
- Inicjacja P^t
- Ocena P^t
- `while not` **warunek stopu** `do`
  - T^t = reprodukcja P^t
  - O^t = krzyżowanie i mutacja T^t (λ osobników)
  - ocena O^t
  - $P^{t+1} = \mu$ najlepszych osobników z O^t
  - $t = t + 1$