

ALGORYTMY EWOLUCYJNE W OPTYMALIZACJI JEDNOKRYTERIALNEJ

Zalety:

- nie wprowadzają żadnych ograniczeń na sformułowanie problemu optymalizacyjnego. Funkcja celu może być wielowartościowa i nieciągła, obszar dopuszczalny – niespójny itp.
- Można je wykorzystać do rozwiązania każdego problemu optymalizacyjnego, tzn. problemu z ciągłymi, dyskretnymi, całkowitymi i mieszanymi zmiennymi decyzyjnymi.

Ogólne sformułowanie jednokryterialnego problemu optymalizacji:

Znajdź wektor zmiennych decyzyjnych: $\mathbf{x}^* = [x_1^*, x_2^*, \dots, x_n^*]^T$

spełniający:

- K ograniczeń nierównościowych $g_k(\mathbf{x}) \geq 0 \quad k = 1, 2, \dots, K$
- M ograniczeń równościowych $h_m(\mathbf{x}) = 0 \quad m = 1, 2, \dots, M$
- i minimalizujący funkcję celu $f(\mathbf{x})$

$$f(\mathbf{x}^*) = \min. f(\mathbf{x})$$

oraz:

$$g_k(\mathbf{x}^*) \geq 0 \quad k = 1, 2, \dots, K$$

$$h_m(\mathbf{x}^*) = 0 \quad m = 1, 2, \dots, M$$

FUNKCJA CELU I FUNKCJA PRZYSTOSOWANIA

Odwzorowanie funkcji celu na funkcję przystosowania (skalowanie funkcji celu):

Selekcja:

- proporcjonalna: skalowanie jest konieczne.
- turniejowa i rankingowa: skalowanie jest nieistotne.

Metody odwzorowania funkcji celu na funkcję dopasowania w przypadku selekcji proporcjonalnej:

- **Sposób najprostszy:** wykorzystanie dodatniej stałej i określenie funkcji przystosowania w postaci:

$$f'(\mathbf{x}) = C - f(\mathbf{x})$$

gdzie :

$$C > f(\mathbf{x}) \quad \text{dla wszystkich } \mathbf{x},$$

lub

$$\begin{aligned} \text{- w iteracji } t \quad C^t &= \max_{j \in J} \{f^j(\mathbf{x})\} \\ J &\text{ – wielkość populacji} \end{aligned}$$

Przykład:

Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = -2 \quad FP(ch5) = 6 \quad FP(ch6) = -3$$

Wybierając stałą $C^t=8$, przeskalowano te wartości do:

$$FP(ch1) = 16 \quad FP(ch2) = 12 \quad FP(ch3) = 13 \quad FP(ch4) = 6 \quad FP(ch5) = 14 \quad FP(ch6) = 5$$

‘Odległości’ pomiędzy kolejnymi FP nie zmieniły się, ale wszystkie FP osiągnęły wartości dodatnie.

- **Statyczne skalowanie liniowe**

$$f'(\mathbf{x}) = a \cdot f(\mathbf{x}) + b$$

a, b – stałe parametry dla wszystkich iteracji.

- **Dynamiczne skalowanie liniowe**

$$f'_t(\mathbf{x}) = a \cdot f(\mathbf{x}) + b_t$$

a – stały parametr dla wszystkich iteracji,
 b_t – parametr ustalany w każdej iteracji.

Przykład:

Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = 2 \quad FP(ch5) = 6 \quad FP(ch6) = 3$$

Przyjmując współczynniki skalujące $a = 5$ oraz $b = 8$, przeskalowano wartości FP chromosomów do:

$$\begin{array}{lll} FP(ch1) = 40+8=48 & FP(ch2) = 20+8=28 & FP(ch3) = 25+8 = 33 \\ FP(ch4) = 10+8=18 & FP(ch5) = 30+8=38 & FP(ch6) = 15+8 = 23 \end{array}$$

‘Odległości’ pomiędzy kolejnymi FP zmieniły się z 4, 1, 3, 4 i 3 na 20, 5, 15, 20 15.

- **Metoda obciążenia**

$$f_t'(\mathbf{x}) = f(\mathbf{x}) + [\tilde{f}(\mathbf{x}) - c \cdot \sigma]$$

c	– stała
σ	- odchylenie standardowe populacji.
$\tilde{f}(\mathbf{x})$	- wartość średnia dopasowania w populacji.

Przykład: Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = 2 \quad FP(ch5) = 6 \quad FP(ch6) = 3$$

Wartość średnia FP wynosi: $\tilde{f}(x) = \frac{8+4+5+2+6+3}{6} = 4,67$, zaś odchylenie standardowe:

$$\begin{aligned} \sigma &= \sqrt{\frac{(8-4,67)^2 + (4-4,67)^2 + (5-4,67)^2 + (2-4,67)^2 + (6-4,67)^2 + (3-4,67)^2}{6}} = \\ &= \sqrt{\frac{11,09 + 0,45 + 1,79 + 7,13 + 1,77 + 2,79}{6}} = \sqrt{\frac{25,02}{6}} = \sqrt{4,17} = 2,04 \end{aligned}$$

Przyjmując wartość stałej $c = 0,2$, otrzymamy następujące wartości FP kolejnych chromosomów:

$$\begin{aligned} FP(ch1) &= 12,262 & FP(ch2) &= 8,262 & FP(ch3) &= 9,262 \\ FP(ch4) &= 6,262 & FP(ch5) &= 10,262 & FP(ch6) &= 2,262. \end{aligned}$$

- **Skalowanie wykładnicze**

$$f'_t(\mathbf{x}) = f_t^\alpha(\mathbf{x})$$

α - wykładnik potęgi bliski jedności, np. 1.005

- **Skalowanie logarytmiczne**

$$f'_t(\mathbf{x}) = b - \log[f_t(\mathbf{x})]$$

b – stała większa od każdej wartości $\log[f_t(\mathbf{x})]$

Przykład: Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = 2 \quad FP(ch5) = 6 \quad FP(ch6) = 3$$

Skalując wykładniczo te wartości z wykładnikiem potęgi równym 1,005, otrzymamy następujące wartości FP chromosomów aktualnej populacji:

$$FP(ch1) = 8,084 \quad FP(ch2) = 4,028 \quad FP(ch3) = 5,040$$

$$FP(ch4) = 2,007 \quad FP(ch5) = 6,054 \quad FP(ch6) = 3,017.$$

‘Odległości’ pomiędzy kolejnymi FP nieznacznie się zwiększyły.

- **Metoda ruchomej bazy (*windowing*)**

$$f'_t(\mathbf{x}) = f_t(\mathbf{x}) - f_w$$

w - szerokość okna (zwykle 2 – 10).

f_w - najgorsza wartość w ostatnich w iteracjach.

Przykład: Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = 2 \quad FP(ch5) = 6 \quad FP(ch6) = 3$$

Przyjmując np. szerokość okna = 3 , wartość FP najgorszego chromosomu w wynosiła 3.

Zatem skalowane wartości FP w populacji t wynoszą:

$$FP(ch1) = 5 \quad FP(ch2) = 1 \quad FP(ch3) = 2 \quad FP(ch4) = -1 \quad FP(ch5) = 3 \quad FP(ch6) = 0$$

Okazało się, że przy takim skalowaniu wartość chromosomu 4 stała się ujemna.

- **Normalizacja**

Problem maksimum:

$$f'_t(\mathbf{x}) = \frac{f_t(\mathbf{x}) - f_{\min} + \gamma}{f_{\max} - f_{\min} + \gamma}$$

Problem minimum:

$$f'_t(\mathbf{x}) = \frac{f_{\max} - f_t(\mathbf{x}) + \gamma}{f_{\max} - f_{\min} + \gamma}$$

Przykład: Dana jest populacja składająca się z 6 chromosomów.

W pokoleniu t uzyskano następujące wartości funkcji przystosowania:

$$FP(ch1) = 8 \quad FP(ch2) = 4 \quad FP(ch3) = 5 \quad FP(ch4) = 2 \quad FP(ch5) = 6 \quad FP(ch6) = 3$$

Wartość minimalna FP wynosi 3, zaś maksymalna jest równa 8.

Przyjmując parametr $\gamma = 2$, znormalizowane dla problemu maksimum wartości FP chromosomów wynoszą:

$$FP(ch1) = \frac{0 + 2}{5 + 2} = 0,286 \quad FP(ch2) = \frac{4 + 2}{5 + 2} = 0,857 \quad FP(ch3) = 1,000$$

$$FP(ch4) = \frac{2 + 2}{5 + 2} = 0,571 \quad FP(ch5) = \frac{6 + 2}{5 + 2} = 0,571 \quad FP(ch6) = 0,714$$

i mieszczą się w przedziale $[0,1]$.

Selekcja turniejowa i rankingowa:

- skalowanie jest nieistotne.

UWZGLĘDNIANIE OGRANICZEŃ

Znajdź wektor: \mathbf{x}^*
minimalizujący funkcję celu: $f(\mathbf{x})$

oraz spełniający

K ograniczeń nierównościowych: $g_k(\mathbf{x}) \geq 0;$

M ograniczeń równościowych: $h_m(\mathbf{x}) = 0.$

Operacje przeprowadzane losowo na chromosomach mogą generować **rozwiązania niedopuszczalne** (tzn. leżące poza obszarem dopuszczalnym wyznaczonym przez ograniczenia).

Metody generujące rozwiązania w obszarze dopuszczalnym:

- **Metoda odrzucania**

W każdej iteracji chromosomy zawierające rozwiązania niedopuszczalne są odrzucane.

(Działa skutecznie gdy obszar dopuszczalny zajmuje większa część całkowitej przestrzeni przeszukiwań.)

- **Metoda poprawiania**

W każdej iteracji chromosomy zawierające rozwiązania niedopuszczalne są „reperowane”, tak aby zawierały rozwiązania z obszaru dopuszczalnego.

(Strategia reperowania zależy od natury problemu i często jest bardziej skomplikowana niż rozwiązywany problem)

- **Metoda funkcji kary**

Przeniesienie na grunt algorytmów ewolucyjnych techniki stosowanej w ‘konwencjonalnych’ metodach optymalizacyjnych (*rozwiązania spoza obszaru dopuszczalnego są ‘karane’ przy wykorzystaniu współczynnika kary*).

Zasada generalna: problem z ograniczeniami jest transformowany do problemu bez ograniczeń, w wyniku włączenia ograniczeń do rozszerzonej funkcji celu:

$$\Phi(\mathbf{x}, r) = f(\mathbf{x}) + r \left\{ \sum_{m=1}^M [h_m(\mathbf{x})]^2 + \sum_{k=1}^K G_k [g_k(\mathbf{x})]^2 \right\}$$

r – współczynnik kontrolujący wielkość członu kary

G_k - operator Heaviside = 0 gdy $g_k(\mathbf{x}) \geq 0$ oraz = 1 gdy $g_k(\mathbf{x}) < 0$.

SELEKCJA TURNIEJOWA W PROBLEMACH OPTYMALIZACJI Z OGRANICZENIAMI

Jedna z najbardziej efektywnych metod rozwiązania zadania programowania nieliniowego z ograniczeniami.

Operator selekcji turniejowej:

W selekcji biorą udział dwa chromosomy populacji rodzicielskiej.

- (i) Jeżeli oba chromosomy są z obszaru niedopuszczalnego, to do populacji tymczasowej wybierany jest chromosom położony bliżej obszaru dopuszczalnego (*na podstawie oceny odległości od brzegu obszaru dopuszczalnego*).

Ocena odległości od obszaru dopuszczalnego:

$$\Psi(\mathbf{x}) = \sum_{m=1}^M [h_m(\mathbf{x})]^2 + \sum_{k=1}^K G_k [g_k(\mathbf{x})]^2$$

W obszarze dopuszczalnym $\Psi(\mathbf{x}) = 0$.

W obszarze niedopuszczalnym $\Psi(\mathbf{x}) > 0$.

Dla żadnego z chromosomów nie oblicza się wartości funkcji celu.

- (ii) Jeżeli jeden chromosom jest z obszaru dopuszczalnego, a drugi z obszaru niedopuszczalnego, to do populacji tymczasowej wybierany jest chromosom należący do obszaru dopuszczalnego.

Dla żadnego z chromosomów nie oblicza się wartości funkcji celu.

- (iii) Jeżeli oba chromosomy są z obszaru dopuszczalnego, to dla obu oblicza się wartość funkcji przystosowania i do populacji tymczasowej przechodzi chromosom o lepszej wartości tej funkcji.

ALGORYTMY EWOLUCYJNE W OPTYMALIZACJI WIELOKRYTERIALNEJ

- Większość problemów optymalizacyjnych rozwiązywanych przy wykorzystaniu algorytmów ewolucyjnych to problemy jednokryterialne, tzn. optymalne rozwiązanie poszukiwane jest jedynie z jednego punktu widzenia, którego matematycznym wyrazem jest funkcja celu.
- Dla rozwiązywania problemów wielokryterialnych, tzn. uwzględniających w poszukiwaniu rozwiązania optymalnego równocześnie kilka punktów widzenia, opracowano niewiele metod wykorzystania algorytmów ewolucyjnych

(za: T. Gwiazda, Algorytmy genetyczne – wstęp do teorii”).

- Znalezienie optymalnego rozwiązania w takim przypadku jest dlatego trudne, że wszystkie punkty widzenia najczęściej nie dadzą się zgrupować w postaci jednej funkcji celu (np. poszukuje się maksymalizacji zysku przy równoczesnej minimalizacji ryzyka).

- Jedną z najbardziej prostych i równocześnie praktycznie stosowanych jest tzw. metoda Schaffera.
- Przy założeniu, że kryteria optymalizacyjne w modelu matematycznym rozwiązywanego problemu nie dadzą się przedstawić w matematycznej postaci jednej funkcji celu należy skorzystać z pojęcia rozwiązania optymalnego w **sensie rozwiązania Pareto**.

Optimum w sensie Pareto – alokacja czynników produkcji lub dóbr konsumpcyjnych, przy której nie można zwiększyć produkcji jednego dobra (lub konsumpcji jednego konsumenta) bez zmniejszenia produkcji innego dobra (lub konsumpcji innego konsumenta). Nazwa terminu pochodzi od nazwiska włoskiego ekonomisty **Vilfreda Pareta**.

(https://pl.wikipedia.org/wiki/Optimum_w_sensie_Pareto)

Lub inaczej:

Rozwiązanie optymalne w sensie Pareto oznacza rozwiązania najlepsze z punktu widzenia wszystkich funkcji celu. Jest to takie rozwiązania, że nie istnieje żadne inne rozwiązanie lepsze chociaż dla jednej funkcji celu, przy pozostałych funkcjach przynajmniej zachowujących swoje wartości. Innymi słowy - jest to takie rozwiązanie (tak dobre rozwiązanie), że żeby je polepszyć (tzn. znaleźć inne lepsze rozwiązanie) dla jednej funkcji celu to tylko kosztem pogorszenia dla innej funkcji.

(http://kbo.ue.poznan.pl/koralewski/programowanie_wkryterialne.htm)

PROBLEM OPTYMALIZACYJNY W SENSIE PARETO

Przyjmijmy, że dane są dwa wektory (np. dwa chromosomy z wartościami zmiennych decyzyjnych kodowanymi zmiennie-przecinkowo):

$$\mathbf{X} = \{x_i\} \quad i = 1, 2, 3, \dots, n \quad \text{oraz} \quad \mathbf{Y} = \{y_i\} \quad i = 1, 2, 3, \dots, n$$

Powiadamy, że wektor \mathbf{X} jest częściowo mniejszy od wektora \mathbf{Y} , jeżeli prawdziwy jest następujący warunek:

$$\bigwedge_{i \in (1:n)} (x_i \leq y_i) \wedge \bigvee_{i \in (1:n)} (x_i < y_i),$$

co zapisujemy jako $\mathbf{X} <_p \mathbf{Y}$ i powiadamy, że wektor \mathbf{X} dominuje nad wektorem \mathbf{Y} .

Jeżeli teraz wektor \mathbf{X} nie jest zdominowany przez żaden inny wektor należący do tego samego zbioru co \mathbf{X} , to nazywamy go **wektorem niezdominowanym**.

Zbiór wszystkich wektorów niezdominowanych tworzy zbiór wektorów (rozwiązań) optymalnych dla rozwiązania zadania optymalizacji wielokryterialnej w sensie Pareto.

Wybór konkretnego rozwiązania z otrzymanego w ten sposób zbioru rozwiązań jest każdorazowo decyzją partykularną (preferującą czyjeś, a nie wszystkich, interesy).

Metoda Schaffera ewolucyjnej optymalizacji wielokryterialnej (1985)

1. Cała populacja P^t chromosomów jest dzielona na rozmiarowo równe subpopulacje, odpowiadające poszczególnym kryteriów optymalizacji (funkcjom celu dla każdego kryterium).
2. Selekcja chromosomów jest przeprowadzana niezależnie w każdej subpopulacji.
3. Wybrane chromosomy rodzicielskie tworzą jedną populację tymczasową T^t , która jest poddana procesowi reprodukcji, tj. mutacji i krzyżowaniu, tworząc w ten sposób następną populację P^{t+1} .

Wadą tej metody jest fakt, że niezależna selekcja najlepszych chromosomów w subpopulacjach z uwagi na pojedyncze kryteria (poszczególne funkcje celu) może prowadzić do eliminacji chromosomów ‘średnich’, tj. dobrych względem wszystkich kryteriów lecz nie najlepszych względem któregoś z nich.

W rezultacie tak otrzymany zbiór rozwiązań optymalnych w sensie Pareto był pozbawiony części swoich rozwiązań.

Usunięcie tej wady zaproponował Goldberg (1989 r.), proponując selekcję rankingową realizowaną w następujący sposób:

1. Cała populacja P^t chromosomów jest dzielona na rozmiarowo równe subpopulacje, odpowiadające poszczególnym kryteriów optymalizacji (funkcjom celu dla każdego kryterium).
 1. Z całej populacji wybierane są wszystkie niezdominowane chromosomy i otrzymują one miejsce '1' w rankingu.
 2. Z pozostałych chromosomów całej populacji wybiera się ponownie wszystkie chromosomy niezdominowane i otrzymują one miejsce '2' w rankingu.
3. Proces ten jest powtarzany aż do wyczerpania całej populacji.
4. Następnie następuje proporcjonalna selekcja kandydatów na rodziców następnej populacji.

