

JOLANTA SOCAŁA (Racibórz)
WITOLD KOSIŃSKI (Warszawa i Bydgoszcz)

Zastosowanie metody funkcji dolnej do badania zbieżności algorytmów genetycznych

Pamięci Profesora Andrzeja Lasoty.

Streszczenie. W badaniu wielu zjawisk przyrodniczych istotną rolę odgrywają operatory Markowa, nieujemne operatory liniowe oraz ich półgrupy. W szczególności rozważana jest asymptotyczna stabilność. A. Lasota i J. A. Yorke w 1982 r. udowodnili, że warunkiem wystarczającym i koniecznym asymptotycznej stabilności dla operatora Markowa jest istnienie nietrywialnej funkcji dolnej. W niniejszej pracy pokazujemy zastosowanie metody funkcji dolnej do badania zachowania algorytmów genetycznych. Rozpatrywane w pracy algorytmy genetyczne, używane do rozwiązywania niegładkich problemów optymalizacyjnych, są wynikiem złożenia dwóch operatorów losowych: selekcji i mutacji. Złożenie tych operacji jest macierzą Markowa.

Słowa kluczowe: operator Markowa, wykładnicza stacjonarność, funkcja dolna, algorytm genetyczny, mutacja, selekcja.

1. Wstęp. W badaniu wielu zjawisk przyrodniczych i procesów technologicznych istotną rolę odgrywają operatory Markowa, nieujemne operatory liniowe, ich półgrupy oraz budowane za ich pomocą modele. Badanie różnych własności wspomnianych operatorów i półgrup pozwala nam wnioskować o zachowaniu matematycznych modeli opisujących zjawiska czy procesy. Jednym z podstawowych problemów jest tutaj badanie różnego typu zbieżności iteracji operatorów bądź półgrup operatorów.

W szczególności, w przypadku operatorów Markowa na przestrzeni funkcji całkowalnych, rozważana jest asymptotyczna stabilność. Jest ona definiowana jako silna zbieżność ciągu kolejnych iteracji operatora Markowa P na dowolnej gęstości f (tzn. zbieżność ciągu $P^n f$) do pewnej ustalonej gęstości f^* , niezależnej od początkowej gęstości f , przy czym gęstość f^* jest jedyną gęstością niezmienniczą operatora P . Przez gęstość rozumiemy tutaj funkcję całkowalną, nieujemną, o całce równej jeden.

A. Lasota i J. A. Yorke [9] udowodnili, że warunkiem wystarczającym i koniecznym asymptotycznej stabilności dla operatora Markowa jest istnienie nietrywialnej *funkcji dolnej*, tzn. takiej nieujemnej funkcji całkowalnej h , że (mówiąc skrótowo) $P^n f \geq h$ dla dowolnej gęstości f i dużych n . Główną różnicą pomiędzy tym twierdzeniem, a klasycznymi uogólnieniami twierdzenia Kreina–Rutmana [1], [23] oraz [14] jest to, że nie zakłada się tutaj żadnego rodzaju zwartości operatora. Twierdzenie to pozostaje prawdziwe nawet w przypadku operatorów przekształcających kulę jednostkową na siebie.

Twierdzenie Lasoty–Yorke’a dało początek całej serii twierdzeń o funkcji dolnej. R. Rudnicki [16] udowodnił twierdzenie o funkcji dolnej dla operatorów dodatnich (bez założenia o zachowaniu całki) na przestrzeni funkcji ciągłych. Wykazał on, że istnienie dodatniej stałej α , spełniającej warunek $P^n f \geq \alpha$ dla dowolnej dodatniej funkcji f oraz dużych n , implikuje wykładniczą stacjonarność operatora P .

Operator nazywamy *wykładniczo stacjonarnym*, jeśli istnieje pewna stała dodatnia λ , pewna dodatnia funkcja f^* oraz ciągły liniowy funkcjonal L taki, że ciąg $\lambda^{-n} P^n f$ jest zbieżny do $f^* L f$ dla dowolnej funkcji dodatniej f . Pojęcie to jest bezpośrednim uogólnieniem na operatory niezachowujące całki pojęcia asymptotycznej stabilności.

A. Zalewska [24] rozpatrzyła przypadek, gdy operator P jest nieujemną kontrakcją na przestrzeni funkcji całkowalnych. A. Lasota i J. A. Yorke zbadali asymptotyczną stabilność dla operatorów Markowa na rodzinie skończonych miar Borelowskich [10]. Ciągi kolejnych iteracji oraz półgrupy operatorów dodatnich rozpatrywali A. Lasota i R. Rudnicki [12] oraz A. Lasota i J. A. Yorke [11].

Współautorka badała przypadek operatorów nieujemnych na przestrzeniach Banacha [18, 19, 20]. Twierdzenia o funkcji dolnej są wygodnymi narzędziami badania zbieżności operatorów Markowa oraz operatorów nieujemnych. Kryteria te były z powodzeniem stosowane między innymi w przypadku operatora Frobeniusa – Perrona transformacji punktowych, dla operatorów całkowych, dla półgrup generowanych przez równania różniczkowe.

Metoda funkcji dolnej pozwoliła zbadać zachowanie różnych modeli fizycznych, biologicznych i medycznych. W niniejszej pracy pokazujemy zastosowanie metody funkcji dolnej do badania zachowania algorytmów genetycznych.

2. Algorytmy genetyczne. Wśród metod używanych do poszukiwania rozwiązań niegładkich problemów optymalizacyjnych od kilku dekad poczesne miejsce zajmują *algorytmy genetyczne*. Algorytmy genetyczne są usytuowane w nurcie *obliczeń ewolucyjnych* i pojawiły się jako pewne implementacje procesów ewolucji doboru organizmów żywych zauważonych w przyro-

dzie, gdy osobniki danego gatunku, charakteryzujące się lepszym przystosowaniem do warunków, w których żyją, mają większe prawdopodobieństwo przeżycia niż osobniki gorzej przystosowane.

Zaobserwowana w naturze ewolucja gatunków, zachodząca w jej trakcie przypadki losowej mutacji kodów genetycznych i jego przenoszenie z rodziców na potomstwo, były inspiracją dla Holland [2] do zaproponowania właśnie takiej nazwy do znanej już wcześniej w literaturze klasy metod rozwiązywania zagadnień optymalizacyjnych (por. Hollstien [3]). Obliczenia bazujące na algorytmach genetycznych wchodzi w grupę tzw. obliczeń ewolucyjnych. Te z kolei należą do rozwijającej się dziedziny wiedzy zwanej inteligencją obliczeniową (ang. *computational intelligence*), a wyrosłej z nurtu badawczego zwanego sztuczną inteligencją. Rozwijane w tej dziedzinie tzw. miękkie metody obliczeniowe (ang. *soft computing*) czasami nie zawsze są dobrze umotywowane i brak im pełnych dowodów poprawności.

Algorytmy genetyczne przeszukują przestrzeń potencjalnych rozwiązań (w języku algorytmów genetycznych - osobników) problemu optymalizacyjnego w sposób losowy, ale nakierowany, bazują też na pewnych heurystykach. Przeszukiwanie nakierowane oznacza istnienie pewnych 'preferencji', które wynikają za każdym razem z oceny każdego odwiedzanego czy wytworzonego punktu przestrzeni. Ocena ta w języku genetycznym to pomiar stopnia dostosowania osobnika do otoczenia (środowiska). Ten aspekt losowości jest niezmiernie ważny – odróżnia on algorytmy genetyczne od całkowicie przypadkowych metod, np. metody Monte Carlo. Oceny dokonuje się wyliczając wartość tzw. *funkcji przystosowania* (dopasowania), będącej nieujemną funkcją określoną przez sam problem optymalizacyjny i występującą w nim funkcję celu.

Skoro kod (chromosom) opisuje budowę wszystkich żywych organizmów i służy do przechowywania i przekazywania materiału genetycznego, to jego uproszczona reprezentacja w obliczeniach komputerowych tworzy ciąg bitów. W ten sposób algorytmy wykorzystujące mechanizmy ewolucji nazywają się *binarnymi algorytmami genetycznymi* (BAG). Były one tematem artykułu Kiesia i Michalewicza opublikowanego w niniejszym Piśmie przed sześciu laty [6].

Główna idea w procesie ewolucji to iteracyjne tworzenia nowych kandydatów na potencjalne rozwiązania, zwane eksploracją przestrzeni poszukiwań, zwane w języku AG tworzeniem populacji (generacji).

W pracy rozpatrujemy uogólnienie szczególnego typu algorytmów genetycznych, tzw. *prostego algorytmu genetycznego* definiowanego w terminach skończonego multizbioru potencjalnych rozwiązań (osobników danej populacji), na którym są określone operacje: mutacji i selekcji, każda z pewnym prawdopodobieństwem. Złożenie tych operacji, działając na daną populację, tworzy nową populację. Istnienie funkcji przystosowania (dopasowania),

określonej na osobnikach populacji, pozwala powiązać prawdopodobieństwo selekcji osobników do nowej populacji, z wartościami, jakie funkcja przystosowania przyjmuje dla danego osobnika. Złożenie tych operacji spełnia własność Markowa, przez co otrzymuje się macierz Markowa. Zbieżność ciągu operacji jest w tym artykule badana metodami funkcji dolnej. W pracy wprowadzamy tylko podstawowe i niezbędne – dla zrozumienia niniejszych wyników – pojęcia z algorytmów genetycznych, nie wdając się w dokładne wyjaśnienia. Dlatego odsyłamy dociekliwego czytelnika do artykułu Kiesia i Michalewicz [6] oraz przeglądowego Rowe [15] i łatwo dostępnych książek w języku polskim, które na początku wprowadzają od podstaw ideę algorytmów genetycznych, a następnie omawiają bardziej szczegółowo zagadnienia ważne dla tego tematu, a także do poprzedniego artykułu autorów [20]. Najważniejsze z nich to książki Michalewicz [13], Goldberga [4] i Cytowskiego [5]. Bardziej zaawansowane podejście z punktu widzenia pewnych układów dynamicznych przynoszą pozycje Vose [22] oraz polskojęzyczna monografia Schaefera [21]. Większość oznaczeń pochodzi właśnie z pozycji [16, 23].

3. Operatory genetyczne. Dla przestrzeni wektorowej \mathbb{R}^s zdefiniujemy normę $\|\cdot\|$ następująco

$$\|(x_0, \dots, x_{s-1})\| = \sum_{i=0}^{s-1} |x_i| \quad \text{dla } (x_0, \dots, x_{s-1}) \in \mathbb{R}^s.$$

Ponadto zdefiniujemy porządek \leq . Mianowicie dla $(x_0, \dots, x_{s-1}), (y_0, \dots, y_{s-1}) \in \mathbb{R}^s$ mamy

$$(x_0, \dots, x_{s-1}) \leq (y_0, \dots, y_{s-1}) \iff x_i \leq y_i \quad \text{dla } i = 0, \dots, s-1.$$

Przestrzeń wektorowa $(\mathbb{R}^s, \leq, \|\cdot\|)$ z tak zdefiniowaną normą oraz porządkiem jest kratą Banacha. Położmy

$$x^- = \max(0, -x) \quad \text{dla } x \in \mathbb{R}^s,$$

gdzie powyższe maksimum jest oczywiście rozumiane jako maksimum w kratce Banacha

Operator $P : \mathbb{R}^s \rightarrow \mathbb{R}^s$ nazywamy *nieujemnym*, jeśli

$$Px \geq 0 \quad \text{dla } x \in \mathbb{R}^s, \quad x \geq 0.$$

W binarnym algorytmie genetycznym działamy na elementach przestrzeni kodowej, która jest obrazem przestrzeni rozwiązań podległej pewnej operacji kodowania binarnego. Elementy przestrzeni kodowej są tutaj reprezentowane przez binarne chromosomy, które mają tę samą długość l , a więc zbiorem wszystkich chromosomów, czyli przestrzenią kodową Z , jest zbiór $Z = \{z_0, \dots, z_{s-1}\}$, gdzie $s = 2^l$, który oznacza uniwersum genetyczne stanowiące zbiór kodów zwanych genotypami osobników próby losowej. Dla poprawności naszych wyników wystarczy założyć, że Z jest dowolnym zbiorem

elementów. W dalszej części naszego artykułu zamiast element przestrzeni kodowej będziemy używać terminu *genotyp*.

Populację, czyli skończony multizbiór o rozmiarze r , zwanym rozmiarem populacji (ang. *PopSize*), która składa się z pewnej liczby tych samych kopii genotypów (elementów przestrzeni kodowej), utożsamiamy z uporządkowaną s -tką liczb wymiernych, ułamków, przy czym każdy element tej s -tki reprezentuje względną liczbę kopii elementu w populacji do liczby wszystkich elementów multizbioru, tzn.

$$(3.1) \quad p_k = \frac{a_k}{r}, \text{ natomiast } p = (p_0, \dots, p_{s-1}),$$

przy czym

$$(3.2) \quad \sum_{k=0}^{s-1} p_k = 1.$$

Tę s -tkę ułamków p odpowiadającą populacji nazywamy *wektorem populacji*. Czasami dla skrót u samo p będziemy nazywać populacją, co nie powinno prowadzić do niejednoznaczności. Własność (3.2) pozwala traktować poszczególne współrzędne populacji p jako prawdopodobieństwa występowania danego elementu z przestrzeni kodowej w populacji. Tym samym, ze względu na (3.2), p staje się wektorem probabilistycznym. Dalsze uszczegółowienia tego wątku można znaleźć w pozycji [22] oraz artykule przeglądowym [15]. W tej definicji nie dopuszczamy, aby współrzędne wektora p przyjmowały dowolne, nawet niewymierne (rzeczywiste) wartości. Jednak, gdy przyjmiemy, że rozmiar populacji dąży do nieskończoności, zbiór możliwych populacji staje się gęsty w następującym zbiorze określonym przez definicję.

DEFINICJA 3.1. Zbiór wszystkich możliwych populacji

$$\Lambda = \{x \in \mathbb{R}^s : \forall_k x_k \geq 0, \sum_{k=0}^{s-1} x_k = 1\}.$$

3.1. Operator selekcji. Dana jest populacja $p = (p_0, \dots, p_{s-1})$ oraz funkcja przystosowania ⁽¹⁾ $f : Z \rightarrow \mathbb{R}^+$, która działa na uniwersum genetycznym, przyporządkując każdemu genotypowi z_k wartość $f(z_k)$, która w języku AG oznacza stopień przystosowania reprezentującego osobnika do otoczenia.

Zakładając stosowanie selekcji proporcjonalnej (por. [6, 14]) operator selekcji działa na wektorze populacji, wyznaczając prawdopodobieństwo, że

⁽¹⁾ Funkcja ta jest złożeniem funkcji celu z funkcją kodującą oraz – jeśli to konieczne – z innymi funkcjami przekształcającymi funkcję celu do funkcji nieujemnej, dla której poszukiwane jest maksimum.

typowy jej element z_k wystąpi w następnej populacji

$$(3.3) \quad q_k = \frac{f(z_k)p_k}{\bar{f}(p)},$$

gdzie $\bar{f}(p)$ jest średnim przystosowaniem populacji p wyznaczonym przez

$$(3.4) \quad \bar{f}(p) = \sum_{k=0}^{s-1} f(z_k)p_k.$$

To pozwoli na formalne wyznaczenie nowej s -tki (wektora) q składającej się z tych prawdopodobieństw przez zastosowanie macierzy diagonalnej S o wymiarze s , w której na głównej przekątnej występują wartości funkcji przystosowania na kolejnych elementach przestrzeni kodowej, tzn.

$$(3.5) \quad S_{kk} = f(z_k).$$

Pozwala to na konsekwentny zapis w postaci

$$(3.6) \quad q = \mathcal{F}p = \frac{1}{\bar{f}(p)}Sp,$$

który określa rozkład prawdopodobieństwa w następnej populacji po zastosowaniu operatora selekcji. Warto zwrócić uwagę, że mnożenie przez czynnik $1/\bar{f}(p)$ pełni rolę normalizacji wyniku działania operatora liniowego S .

3.2. Operator mutacji. Przejdźmy do operatora mutacji⁽²⁾. Wyjdźmy z dowolnego elementu z_j . Wiemy, że prawdopodobieństwo występowania tego elementu jest równe q_j . Prawdopodobieństwo przejścia w element z_i na drodze mutacji z populacji q jest równe

$$(3.7) \quad \sum_{j=0}^{s-1} U_{ij}q_j,$$

gdzie U_{ij} jest elementem macierzy U opisującej prawdopodobieństwa mutacji z elementu z_j w element z_i w przypadku $i \neq j$. Gdy $i = j$ jest to prawdopodobieństwo przetrwania elementu z_i w trakcie mutacji.

Sposób wyznaczania elementów tej macierzy pokazuje następujący przykład odwołujący się do kodowania binarnego i genotypu reprezentowanego przez skończony ciąg zer i jedynek. Gdy z_i różni się od z_j na c pozycjach, to

$$(3.8) \quad U_{ij} = \mu^c(1 - \mu)^{l-c}.$$

Dla poprawności naszych wyników zawartych w następnych rozdziałach nie ma potrzeby ograniczania się jedynie do macierzy mutacji U , której

⁽²⁾ Dla prostego operatora genetycznego naturalnym jest rozpatrzyć najpierw mutację binarną, równomierną, o parametrze μ . Oznacza to, że dowolny gen w chromosomie, może być zmutowany z prawdopodobieństwem μ .

elementy dane są przez (3.8). Wyniki będą poprawne dla przypadku ogólniejszego, w szczególności dla niebinarnych operatorów mutacji. Jedynie, czego będzie się wymagać to to, aby elementy macierzy U były nieujemne oraz ich suma w każdej kolumnie była równa jeden. To oznacza, że macierz U przeprowadza wektory prawdopodobieństwa w wektory prawdopodobieństwa, a to oznacza, że jest macierzą Markowa [8].

Wyniki naszego artykułu dotyczą algorytmu genetycznego, w którym został pominięty następny element prostego algorytmu genetycznego, a mianowicie krzyżowanie.

Składając operacje mutacji i selekcji otrzymujemy zależność

$$(3.9) \quad p(t+1) = U \circ \mathcal{F}p(t) = \frac{1}{\bar{f}(p(t))} USp(t),$$

gdzie zmienna t oznacza numer populacji, kroku iteracji.

Zdefiniujemy operator G następująco

$$(3.10) \quad G(p) = U \circ \mathcal{F}p = \frac{1}{\bar{f}(p)} USp.$$

Będziemy badać zachowanie ciągów $G^t(p)$ dla dużych $t \in \mathbb{N} = \{0, 1, \dots\}$ i dowolnej populacji początkowej p .

4. Zbieżność algorytmów genetycznych

DEFINICJA 4.1. Będziemy mówili, że model jest wykładniczo stacjonarny, jeśli istnieją $p^* \in \Lambda$, $\lambda \in \mathbb{R}_+$ oraz $L : \mathbb{R}^s \rightarrow \mathbb{R}$ funkcjonal liniowy i nieujemny takie, że

$$(4.1) \quad USp^* = \lambda p^*,$$

$$(4.2) \quad \lim_{t \rightarrow \infty} \|\lambda^{-t} (US)^t p - L(p) \cdot p^*\| = 0 \quad \text{dla } p \in \mathbb{R}^s.$$

Będziemy używać następujących warunków:

(I') Dla dowolnych genotypów z_i oraz z_j prawdopodobieństwo przejścia w skończonej ilości mutacji od z_i do z_j jest dodatnie.

(II') Prawdopodobieństwo tego, że dowolny genotyp z_i nie ulegnie mutacji jest dodatnie.

TWIERDZENIE 4.1. *Załóżmy, że spełnione są warunki (I'), (II'). Wówczas model jest wykładniczo stacjonarny.*

UWAGA 4.1. W Twierdzeniu 4.1. żadnego z warunków (I'), (II') nie można opuścić. Pokazują to Przykłady 4.1, 4.2.

UWAGA 4.2. Z Twierdzenia 4.1. wynika, że jeśli spełnione są warunki (I'), (II'), to dla $p \in \mathbb{R}^s$ mamy

$$(4.3) \quad \lim_{t \rightarrow \infty} \lambda^{-t} \|(US)^t p\| = |L(p)|.$$

Daje nam to informację o ilościowym zachowaniu populacji w przypadku modelu, dla którego nie dokonywalibyśmy unormowania populacji przy każdym kroku. W takim przypadku liczba $\|(US)^t p\|$ (która może być rozumiana jako wielkość populacji) rośnie do ∞ dla $\lambda > 1$, maleje do 0 dla $\lambda < 1$, dąży do $|L(p)|$ dla $\lambda = 1$.

DEFINICJA 4.2. Zdefiniujmy pewien zbiór $\Lambda_+ \subset \mathbb{R}^s$ następująco

$$\Lambda_+ = \{x \in \mathbb{R}^s : \forall_k \quad x_k \geq 0, \quad \|x\| \neq 0\}.$$

WNIOSEK 4.1. Z Twierdzenia 4.1 wynika, że jeśli spełnione są warunki (I'), (II'), to dla $p \in \Lambda_+$ mamy

$$(4.4) \quad \lim_{t \rightarrow \infty} \frac{(US)^t p}{\|(US)^t p\|} = p^*.$$

TWIERDZENIE 4.2. Załóżmy, że dla algorytmu spełnione są warunki (I'), (II'). Wówczas istnieje $p^* \in \Lambda$ takie, że $G(p^*) = p^*$ oraz

$$(4.5) \quad \lim_{t \rightarrow \infty} \|G^t(p) - p^*\| = 0 \quad \text{dla } p \in \Lambda.$$

Tak więc dla operatora G istnieje populacja p^* , która jest jego punktem stałym. Łatwo pokazać, że jest to jedyna populacja o tej własności. Ponadto niezależnie od początkowej populacji $p \in \Lambda$ ciąg kolejnych iteracji $G^t(p)$ jest zbieżny do p^* .

Przykład 4.1. Pokażemy teraz, że w naszych twierdzeniach warunek (I') jest niezbędny. Załóżmy, że $s = 2$, $f(z_k) = 1$ dla $k = 0, 1$ oraz

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Wówczas spełniony jest warunek (II'), a warunek (I') nie jest spełniony. Ustalmy teraz dowolną populację $p \in \Lambda$. Wówczas mamy $US(p) = p$ oraz $G(p) = p$. Tak więc nie ma wykładniczej stacjonarności oraz $G^t(p)$ nie zmierzają do wspólnej granicy niezależnej od p .

Przykład 4.2. Pokażemy teraz, że w naszych twierdzeniach warunek (II') jest niezbędny. Załóżmy że $s = 2$, $f(z_k) = 1$ dla $k = 0, 1$ oraz

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Wówczas spełniony jest warunek (I'), a warunek (II') nie jest spełniony. Dla $p = (1, 0)$ oraz $t = 1, 2, \dots$ mamy

$$(US)^t p = \begin{cases} (1, 0), & \text{jeśli } t \text{ jest liczbą parzystą;} \\ (0, 1), & \text{jeśli } t \text{ jest liczbą nieparzystą.} \end{cases}$$

Tak więc nie ma wykładniczej stacjonarności. Ponieważ, w tym przypadku, $G^t p = (US)^t p$ dla $t \in \mathbb{N}$, więc $G^t(p)$ nie zmierza do żadnego $p^* \in \Lambda$.

5. Dowody twierdzeń. W tej części pracy pokażemy dowody Twierdzeń 4.1, 4.2, przy czym wykorzystamy tutaj metodę funkcji dolnej. Twierdzenie 4.1 można również udowodnić inaczej przy pomocy klasycznego twierdzenia Perrona.

DOWÓD TWIERDZENIA 4.1 Dla $i, j \in \{0, \dots, s - 1\}$ połóżmy

$$(5.1) \quad n(i, j) = \min \left\{ l \in \mathbb{N} : \exists_{\{s_0, \dots, s_l \in \{0, \dots, s-1\}\}} s_0 = i, \quad s_l = j, \right. \\ \left. U_{s_{k+1}s_k} > 0 \text{ dla } k = 0, \dots, s_l - 1 \right\}$$

oraz

$$(5.2) \quad n = \max \{ n(i, j) : i, j \in \{0, \dots, s - 1\} \}.$$

Łatwo pokazać, na podstawie warunku (II'), że wszystkie wyrazy macierzy $(US)^n$ są dodatnie. Oznaczmy $[b_{ij}] = (US)^n$. Niech $b = \inf \{ b_{ij} : i, j \in \{0, \dots, s - 1\} \}$, $c = \sup \{ b_{ij} : i, j \in \{0, \dots, s - 1\} \}$. Zdefiniujmy wektor $g = \frac{b}{c \cdot s} (1, \dots, 1) \in \Lambda_+$. Jak zauważyliśmy wcześniej, przestrzeń wektorowa $(\mathbb{R}^s, \leq, \|\cdot\|)$ ze zdefiniowaną przez nas normą oraz porządkiem jest kratą Banacha. Pokażemy, że spełnione są warunki (I), (II) z Twierdzenia 1 [17], przy czym g pełni rolę funkcji h występującej w Twierdzeniu 1 [17], zbiór Λ_+ pełni rolę D , a operator US pełni rolę operatora P . Najpierw pokażemy, że spełniony jest warunek (I). Ustalmy $p \in \Lambda_+$. Mamy

$$(5.3) \quad ((US)^n p)_k = \sum_{i=0}^{s-1} b_{ki} p_i \geq b \|p\|$$

oraz

$$(5.4) \quad \|(US)^n p\| = \sum_{k=0}^{s-1} \left(\sum_{i=0}^{s-1} b_{ki} p_i \right) \leq c \|p\| \cdot s.$$

Z (5.3) i (5.4) mamy

$$(5.5) \quad \frac{(US)^n p}{\|(US)^n p\|} \geq g \text{ dla } p \in \Lambda_+.$$

Skoro U jest macierzą Markowa, a S jest macierzą diagonalną o wyrazach dodatnich na przekątnej, łatwo pokazać indukcyjnie, że

$$(5.6) \quad \|(US)^t p\| \neq 0 \text{ dla } p \in \Lambda_+, t \in \mathbb{N}.$$

Ustalmy teraz $p \in \Lambda_+$ oraz $t > n$. Z (5.6) mamy $(US)^{t-n} p \in \Lambda_+$. Stąd i na podstawie (5.5)

$$\frac{(US)^t p}{\|(US)^t p\|} = \frac{(US)^n ((US)^{t-n} p)}{\|(US)^n (US)^{t-n} p\|} \geq g.$$

Tak więc

$$\lim_{t \rightarrow \infty} \left\| \left(\frac{(US)^t p}{\|(US)^t p\|} - g \right)^- \right\| = 0 \quad \text{dla } p \in \Lambda_+$$

i warunek (I) jest spełniony. Pokażemy teraz, że zachodzi warunek (II). Ustalmy $p \in \Lambda_+$. Mamy

$$((US)^n g)_k = \sum_{i=0}^{s-1} b_{ki} g_i \geq b \|g\|.$$

Stąd

$$((US)^n p)_k \leq \frac{c \|p\|}{b \|g\|} \cdot b \|g\| \leq \frac{c \|p\|}{b \|g\|} ((US)^n g)_k.$$

Tak więc

$$(US)^n p \leq \frac{c \|p\|}{b \|g\|} (US)^n g.$$

Stąd i z nieujemności operatora US dla $t > n$ mamy

$$(US)^t p = (US)^{t-n} ((US)^n p) \leq (US)^{t-n} \left(\frac{c \|p\|}{b \|g\|} (US)^n g \right) = \frac{c \|p\|}{b \|g\|} (US)^t g.$$

Tak więc

$$\limsup_{t \rightarrow \infty} \frac{\|(US)^t p\|}{\|(US)^t g\|} < \infty \quad \text{dla } p \in \Lambda_+.$$

Na podstawie Twierdzenia 1 [17] model jest wykładniczo stacjonarny. \square

DOWÓD TWIERDZENIA 4.2 Łatwo pokazać, że

$$(5.7) \quad G(p) \in \Lambda \quad \text{dla } p \in \Lambda.$$

Na podstawie wzorów (3.10) i (4.1) mamy

$$G(p^*) = \frac{\lambda}{\bar{f}(p^*)} p^*.$$

Z (5.7) wynika, że $G(p^*) \in \Lambda$, więc $\lambda/\bar{f}(p^*) = 1$. Stąd $G(p^*) = p^*$. Ustalmy teraz $p \in \Lambda$. Pokażemy indukcyjnie, że dla $t \in \mathbb{N}$ zachodzi

$$(5.8) \quad G^t p = \frac{(US)^t p}{\|(US)^t p\|}.$$

Dla $t = 0$ teza wynika z faktu, że $p \in \Lambda$. Załóżmy teraz, że (5.8) zachodzi dla pewnego ustalonego $t \in \mathbb{N}$. Pokażemy że nasz wzór zachodzi dla $t + 1$. Oznaczmy $[\sigma_{ij}] = US$. Mamy oczywiście

$$(5.9) \quad \sum_{k=0}^{s-1} \sigma_{kj} = f(z_j).$$

Zachodzi

$$\begin{aligned}\bar{f}\left(\frac{(US)^t p}{\|(US)^t p\|}\right) &= \sum_{k=0}^{s-1} f(z_k) \left(\frac{(US)^t p}{\|(US)^t p\|}\right)_k = \\ &= \frac{1}{\|(US)^t p\|} \sum_{k=0}^{s-1} f(z_k) ((US)^t p)_k.\end{aligned}$$

Stąd i z (3.10) mamy

$$\begin{aligned}G^{t+1}(p) &= G(G^t(p)) = G\left(\frac{(US)^t p}{\|(US)^t p\|}\right) = \\ &= \frac{1}{\bar{f}\left(\frac{(US)^t p}{\|(US)^t p\|}\right)} US \left(\frac{(US)^t p}{\|(US)^t p\|}\right) = \frac{(US)^{t+1} p}{\sum_{k=0}^{s-1} f(z_k) ((US)^t p)_k}.\end{aligned}$$

Ponadto na podstawie (5.9)

$$\begin{aligned}\frac{(US)^{t+1} p}{\|(US)^{t+1} p\|} &= \frac{(US)^{t+1} p}{\sum_{k=0}^{s-1} ((US)^{t+1} p)_k} = \frac{(US)^{t+1} p}{\sum_{k=0}^{s-1} \sum_{j=0}^{s-1} \sigma_{kj} ((US)^t p)_j} = \\ &= \frac{(US)^{t+1} p}{\sum_{j=0}^{s-1} ((US)^t p)_j \sum_{k=0}^{s-1} \sigma_{kj}} = \frac{(US)^{t+1} p}{\sum_{j=0}^{s-1} ((US)^t p)_j f(z_j)}.\end{aligned}$$

Tak więc

$$G^{t+1} p = \frac{(US)^{t+1} p}{\|(US)^{t+1} p\|}.$$

Teza twierdzenia wynika z (4.4). \square

Artykuł powstał we współpracy, w której drugi autor (W.K.) był wspierany projektem KBN Nr 3 T11 C007 28. Autorzy są wdzięczni Stefanowi Kotowskiemu za cenne dyskusje.

Bibliografia

- [1] H. Amann, *Fixed point theorems and nonlinear eigenvalue problems*, SIAM Rev. 18 (1976), 620–709.
- [2] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.
- [3] R. B. Hollstien, *Artificial Genetic Adaptation in Computer Control Systems*, Ph.D. Thesis, University of Michigan, 1971.
- [4] D. E. Goldberg, *Algorytmy genetyczne i ich zastosowania*, WNT, Warszawa, 1995.
- [5] J. Cytowski, *Algorytmy genetyczne: podstawy i zastosowania*, Seria: Problemy Współczesnej Nauki - Teoria i Zastosowania Nr 18, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1996.
- [6] P. Kieś i Z. Michalewicz, *Podstawy algorytmów genetycznych*, Matematyka Stosowana. Matematyka dla Społeczeństwa, 1 (44), 2000, 68–91.
- [7] W. Kosiński, S. Kotowski, J. Socała, *On Asymptotic Behaviour of a Binary Genetic Algorithm*. Annales UMCS, Informatica AI, 4, 2006, 180–188, Proceedings of the

Scientific Session organized during XXIst Fall Meeting of the Polish Information Processing Society.

- [8] A. Lasota, *Asymptotyczne własności półgrup operatorów Markowa*, Matematyka Stosowana. Matematyka dla Społeczeństwa, **3** (45), 2002, 39–51.
- [9] A. Lasota, J. A. Yorke, *Exact dynamical systems and the Frobenius–Perron operator*, Trans. Amer. Math. Soc. 273 (1982), 375–384.
- [10] A. Lasota, J. A. Yorke, *Lower bound technique for Markov operators and iterated function systems*, Random Comput. Dynam. 2 (1994), 41–77.
- [11] A. Lasota, J. A. Yorke, *When the long-time behaviour is independent of the initial density*, SIAM J. Math. Anal. 27 (1996), 221–240.
- [12] A. Lasota, R. Rudnicki, *Asymptotic behaviour of semigroups of positive operators on $C(X)$* , Bull. Pol. Ac. Sci.: Math. 36 (1988), 151–159.
- [13] Z. Michalewicz, *Algorytmy genetyczne + struktury danych = programy ewolucyjne*, WNT, Warszawa, 1996.
- [14] R. D. Nussbaum, *Eigenvectors of nonlinear positive operators and the linear Krein–Rutman theorem*, Fixed Point Theory, Proc. Conf. Sherbrooke, Lecture Notes Math., Springer 886 (1980), 309–330.
- [15] J. E. Rowe, *The dynamical system models of the simple genetic algorithm*, w Theoretical Aspects of Evolutionary Computing, Leila Kallel, Bart Naudts, Alex Rogers (Eds.), Springer, 2001, pp. 31–57.
- [16] R. Rudnicki, *Asymptotic properties of the iterates of positive operators on $C(X)$* , Bull. Pol. Ac. Sci.: Math. 34 (1986), 181–187.
- [17] J. Socała, *Asymptotic behaviour of the iterates of nonnegative operators on a Banach lattice*, Ann. Polon. Math 68 (1) (1998), 1–16.
- [18] J. Socała, *Asymptotic behaviour of the iterates of nonnegative operators on a Banach spaces with a cone*, Bull. Pol. Ac. Sci.: Math. 50 (2) (2002), 179–187.
- [19] J. Socała, *Asymptotic behaviour of semigroups of nonnegative operators on a Banach lattice*, Ann. Polon. Math 82 (2) (2003), 95–103.
- [20] J. Socała, W. Kosiński, S. Kotowski, *O asymptotycznym zachowaniu prostego algorytmu genetycznego*, Matematyka Stosowana. Matematyka dla Społeczeństwa, PTM, Warszawa, **6** (47), 2005, 70–86.
- [21] R. Schaefer, *Podstawy genetycznej optymalizacji globalnej*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2002.
- [22] M. D. Vose, *The Simple Genetic Algorithm: Foundation and Theory*, MIT Press, Cambridge, MA, 1999.
- [23] P. P. Zabreiko, M. A. Krasnosel’skiĭ, Yu. V. Pokornyi, *A certain class of positive linear operators*, Funktsional. Anal. i Prilozhen. 5 (4) (1971), 9–17 (in Russian).
- [24] A. Zalewska-Mitura, *A generalization of the lower bound function theorem for Markov operators*, Univ. Jagell. Acta Math. 31 (1994), 79–85.

Jolanta Socała

Instytut Techniki i Matematyki

Państwowa Wyższa Szkoła Zawodowa w Raciborzu

ul. Słowackiego 55, 47-400 Racibórz

E-mail: jolanta.socala@pwsz.raciborz.edu.pl

Witold Kosiński
Polsko-Japońska Wyższa Szkoła Technik Komputerowych
Katedra Systemów Inteligentnych
ul. Koszykowa 86, 02-008 Warszawa

Uniwersytet Kazimierza Wielkiego
Instytut Mechaniki Środowiska i Informatyki Stosowanej
ul. Chodkiewicza 30, 85-064 Bydgoszcz
E-mail: wkos@pjawstk.edu.pl

**Lower-bound function method in the convergence analysis
of genetic algorithms**

Abstract. Markovian operators, non-negative linear operators and its subgroups play a significant role for the description of phenomena observed in the nature. Research on asymptotic stability is one of the main issues in this respect. A. Lasota and J. A. Yorke proved in 1982 that the necessary and sufficient condition of the asymptotic stability of a Markovian operator is the existence of a non-trivial lower-bound function. In the present paper it is shown how the method of lower-bound function can be applied to the investigation of genetic algorithms. Genetic algorithms considered used for solving of non-smooth optimization problems are compositions of two random operators: selection and mutation. The compositions are Markovian matrices.

Key words: Markov operator, exponential stationarity, lower-bound function, genetic algorithm, mutation, selection.

(wplynęło 3 czerwca 2007 r.)