

Model ewolucyjny

Michał Stanowski, Stanisław Gołębiewski

April 15, 2024

1 Wprowadzenie

Celem zadania było zaimplementowanie w języku Python modelu ewolucji opartego na Geometrycznym Modelu Fishera (GMF). Rozród osobników połączono z dostosowaniem do środowiska, opierając się na funkcji gęstości rozkładu normalnego. Dostosowanie osobników zależy od ich odległości od optymalnego genotypu, który podlega (i) liniowym zmianom związanym z globalnym ociepleniem, (ii) nagłym, gwałtownym zmianom związanym z uderzeniem meteorytu. Nasz program umożliwia śledzenie w czasie rzeczywistym ewolucji populacji, a także jej liczebności, zmienności i odległości od optimum najlepiej dostosowanego osobnika. **Dodatkowo, po** zakończonej symulacji możliwe jest prześledzenie filogenezy osobników z populacji założycielskiej oraz historii przodków wybranego osobnika z ostatniej generacji. Zaimplementowano także wyświetlanie prostych wykresów liczebności i zmienności populacji. Możliwe jest także porównanie średniej wartości fenotypów w populacji oraz wartości najlepiej przystosowanego osobnika, aby zobaczyć, które geny decydowały o dostosowaniu w trakcie trwania symulacji. Wreszcie nasz program umożliwia prowadzenie symulacji dla wielowymiarowych genomów, ponieważ korzystamy z metody PCA przy wizualizacji.

Program ma swoje wady - populacja wysyca pojemność środowiskową bardzo szybko, a mechanizm odcięcia (choć pozwala na przeżycie także osobników mniej dostosowanych) jest brutalny i wydaje się autorom nienaturalny. Dodatkowo, wizualizacja działa poprawnie dla małych wartości σ - dla większych prawa plansza jest zbyt mała, ale uważamy, że jest to nieistotny detal, sama symulacja działa w sposób poprawny.

2 Główna pętla programu

Główna pętla programu składa się z czterech zasadniczych części - na początku **zmieniamy optimum**. Zmiana optimum może przyjąć formę globalnego ocieplenia i jest to zmiana liniowa lub formę uderzenia meteorytu, który zmienia optimum w sposób znaczący. Globalne ocieplenie działa w jednym kierunku, tzn. wartości są wyłącznie dodawane do genotypu optymalnego, zaś uderzenie meteorytu może spowodować zarówno zwiększenie wartości optymalnych, jak i ich zmniejszenie. Toteż uderzenie meteorytu może doprowadzić do sytuacji, że populacja, która ledwo trzyma się przy życiu, po uderzeniu może stać się świetnie dostosowana ze względu na *cofnięcie się* optimum. Następnie

wykonywane są **mutacje i liczone jest dostosowanie**. Mutacje dotyczą w różnym stopniu miejsc zmiennych (pierwsza połowa genomu) i miejsc konserwowanych (druga). Mutacje, podobnie jak w przypadku optimum działają w sposób liniowy. Dostosowanie jest liczone w oparciu o funkcję gęstości rozkładu normalnego. W trzecim kroku następuje **rozmnażanie** oparte o dostosowanie osobników. Wyróżniamy $max_num_children + 1$ klas rozrodczych, do których wpadają osobniki w oparciu o dostosowanie. Po rozrodzie *obcinamy* populację do maksymalnej ilości osobników za pomocą funkcji, która przypisuje malejące prawdopodobieństwa znalezienia się w następnej generacji wszystkim osobnikom, co powoduje, że także osobniki mniej dostosowane mogą przeżyć do następnego kroku.

3 Filogeneza

W części filogenetycznej programu można prześledzić filogenezę populacji ancestralnej. Nasz program umożliwia zapisanie sieci do pliku, który następnie można wizualizować np. w programie *Cytoscape*. Każdy wierzchołek sieci ma konstrukcję **generacja.ranking**, gdzie ranking rozumiemy jako pozycję w liście posortowanej po malejącym dostosowaniu. Przykładowa sieć została przedstawiona na Fig. 1.

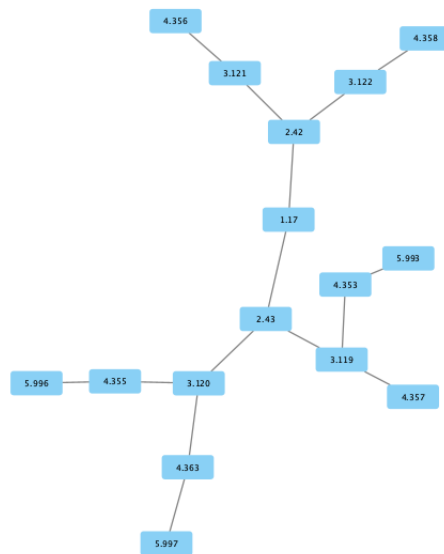


Figure 1: Sieć wygenerowana za pomocą programu (wizualizacja: Cytoscape)

Należy podkreślić, że ręczna analiza sieci filogenetycznych jest możliwa praktycznie wyłącznie dla małych sieci - osobniki bardzo dobrze dostosowane cechują się ogromnymi sieciami, których analiza możliwa jest wyłącznie informatycznie.

W części filogenetycznej możliwa jest także analiza dostosowania przodków wybranego osobnika pochodzącego z ostatniej generacji. Na wykresie skrótem **AR** oznaczono ranking osobnika z populacji założycielskiej, od którego pochodzi wybrany osobnik. Przykładowy

wykres przedstawiono na Fig. 2.

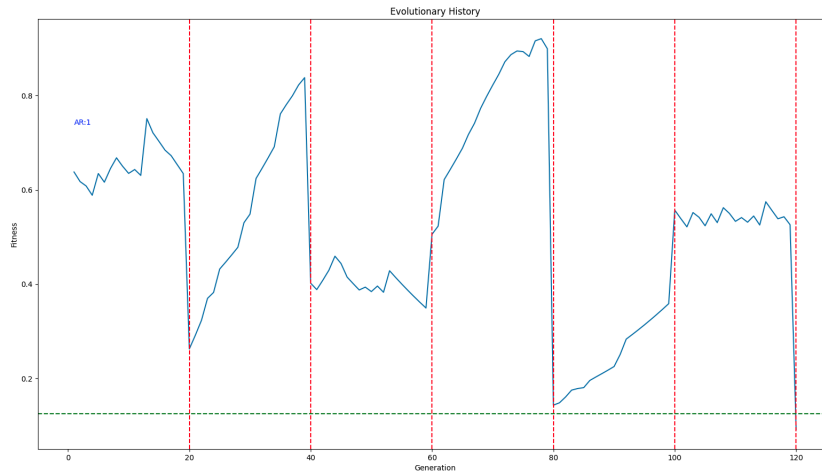


Figure 2: Wykres dostosowania przodków wybranego osobnika ostatniej generacji

4 Wykresy

Poniżej przedstawiono podstawowe wykresy możliwe do wygenerowania w naszym programie. Na Fig. 3 przedstawiono podstawowy wykres liczebności, a na Fig. 4 wykres zmienności (zmienność jest liczona jako suma odchyłeń standardowych każdego genu w populacji). Na Fig. 5 przedstawiono natomiast wykres zmian wartości poszczególnych genów dla populacji (GP, średnia) oraz najbardziej dostosowanego osobnika (GF).

5 Kod

Na kod składają się 4 pliki, które obsługują środowisko, osobnika, pętlę główną oraz wizualizację. Łącznie pliki te mają ok. 700 linijek, przy czym znaczna większość przypada na plik z wizualizacją. Biblioteką, z której korzystano najczęściej była pygame. Do kwestii okołomatematyczno-statystycznych korzystano z matplotlib, numpy, scipy, scikit oraz networkx.

6 Podsumowanie

Uważamy, że nasz program w bardzo przyjemny dla oka sposób umożliwia prześledzenie prostego modelu ewolucji wraz z odtworzeniem filogenezy wybranych osobników z populacji ancestralnej. W przyszłości należałoby jednak popracować nad wizualizacją, która dla dużych wartości σ działa w sposób bardzo nieefektywny oraz nad selekcją po urodzeniu młodych osobników, która w naszym mniemaniu jest zbyt brutalna.

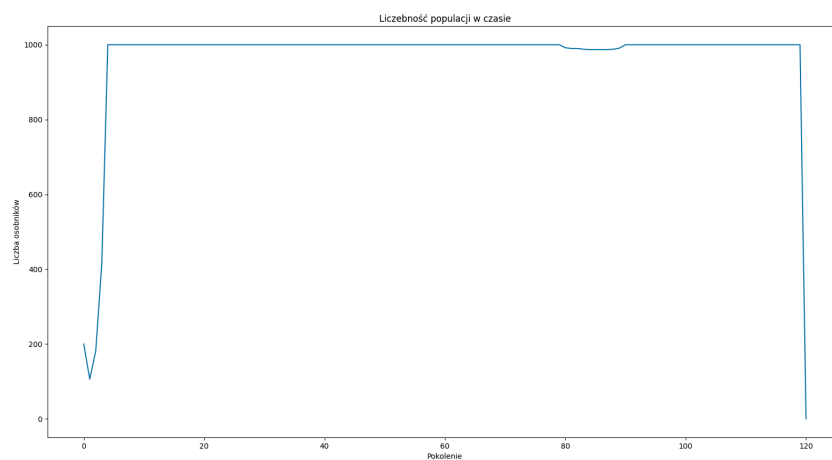


Figure 3: Wykres zmian liczebności populacji w czasie

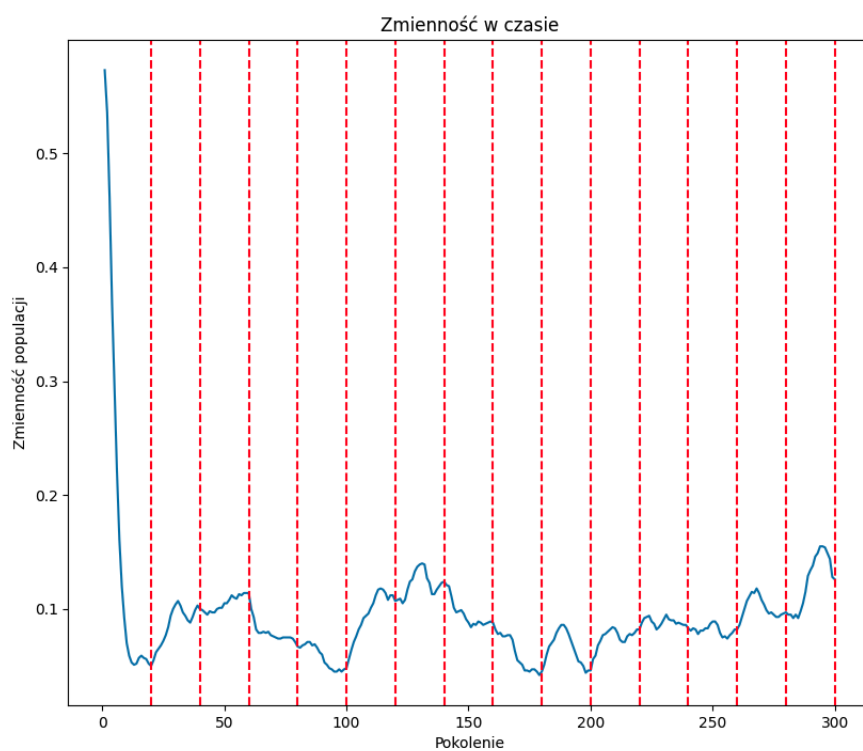


Figure 4: Wykres zmian zmienności populacji w czasie



Figure 5: Wykres zmian wartości genów dla populacji (GP) i osobnika najlepiej dostosowanego (GF) w czasie