Symulacje Komputerowe
Sprawozdanie 1

1 Ogólne uwagi

Sprawozdania można wykonywać w grupach dwu- lub trzyosobowych. Prace indywidualne są akceptowane, ale ze względu na ilość pracy doradza się połączenie sił. Ocena za kody, staranność, poprawność, wnioski i prezentację jest przyznawana z równą wagą za każde zadanie oddzielnie. Jedynie terminowość liczy się oddzielnie.

Rozwiązanie każdego zadania powinno się składać ze wstępu teoretycznego, sformułowania problemu badawczego, propozycji rozwiązania problemu (tutaj znaleźć się powinny np. ważniejsze pseudokody), przedstawieniu wyników (np. odpowiednio dobrane wykresy i tabele) i wniosków, w których można podsumować wyniki, postawić ewentualne hipotezy badawcze, postawić pytania odnośnie ewentualnych kierunków dalszych prac. Do sprawozdania załączamy listę kodów. Kody do każdego zadania powinny się znaleźć w osobnym pliku .ipynb lub .m. W przypadku korzystania z Markdowna możliwe jest umieszczenie kodów w pliku ze sprawozdaniem, koniecznie pod rozwiązaniem danego zadania. Kody w kolejnych komórkach w Jupyterze powinny być modularne, tzn. deklaracje osobnych funkcji powinny być umieszczone w osobnych komórkach, podobnie wywołania dla różnych analiz też powinny się znaleźć w osobnych komórkach. Cała praca (kody+sprawozdanie) powinna zostać wysłana w formacie .rar i podpisana nazwiskami wszystkich członków grupy w następujący sposób: KaczyńskiTusk1.rar.

W sprawozdaniu rozwiązania zadań powinny być samodzielne: nie korzystamy z kodów innych grup, można jednak inspirować się cudzymi ideami, czy wspólnie opracowywać pseudokody, dowody twierdzeń itp. (w takim wypadku należy napisać, do kogo należał pomysł).

Zadania powinny działać pod Linuksem (w przeciwnym wypadku -15%).

Wskazówki co do pisania sprawozdań:

- Zbiór esejów nt. wizualizacji danych http://www.biecek.pl/Eseje/
- Ogólne uwagi ze strony (+linki) dra Płociniczaka: `http://prac.im.pwr.edu.pl/~plociniczak/doku.php?id=dyplomanci`
- Wskazówki ze strony pana Ślęzaka: `http://prac.im.pwr.wroc.pl/~slezak/rapWytyczne.pdf`
- *Czysty kod*, R. C. Martin

W sprawie pytań lub wątpliwości można pisać maile, łapać w Instytucie, w kinie, na Facebooku, w górach, gdzie bądź. Postaram się służyć pomocą. Warto konsultować postępy prac na bieżąco.

2 Pardwy i Wilcy

Motto zadania: „Potrzeba bardzo wiele sów, aby nie powiedzieć pardwy” (via Facebook, „Tylko pardwa jest ciekawa”)

W nieskończonym lesie żyje $m = (m_{♂}, m_{♀})$ pardw i $n = (n_{♂}, n_{♀})$ wilków. Na początku pardwy i wilcy są rozmieszczeni losowo (jednostajnie) w kole o promieniu R i środku w $(0,0)$ (1 pkt.).

Gdyby istniał tylko jeden wilk lub tylko jedna pardwa, przemieszczałyby się one losowo, a ich trajektorie przypominałyby ruch Browna $B_w((0,0), \sigma_w \cdot I_2)$, $B_p((0,0), \sigma_p \cdot I^2)$, gdzie I_2 jest macierzą jednostkową rozmiaru 2 (2 pkt.). Wilki i pardwy są jednak interaktywne, przez co ich rzeczywiste trajektorie to ruch Browna zmodyfikowany o dryf zależny od położeń pozostałych osobników. Odległość między dwoma osobnikami a i b oznaczamy przez $\rho(a,b)$, natomiast kierunek wektora (a,b) przez $\vec{ab}$ (dla skrócenia notacji przyjmujemy zawsze, że długość tego wektora zawsze wynosi 1).

Pardwy:

1. Dorosłe samce trzymają się od siebie z daleka. Jeśli dwa p_1, p_2 samce znajdą się od siebie w odległości mniejszej niż $h_{p,♂,♂}$, to do dryfu należy dodać wektor $\frac{-3}{\rho(p_1,p_2)} p_1 \vec{p}_2$ (tzn. w każdej jednostce czasu jeden samiec oddala się od drugiego z taką prędkością) do czasu, aż oddalą się od siebie ponownie na odległość $h_{p,♂,♂}$.

2. Samice są wabione przez głos samców z odległości $h_{p,♂,♀}$. Jeśli samica p_1 znajdzie się w odległości mniejszej niż $h_{p,♂,♀}$ od samca p_2, to do jej dryfu należy dodać $p_1 \vec{p}_2$. Jeśli samiec znajdzie się w odległości mniejszej niż $h_{p,♂,♀}/3$ od samicy, to do jego dryfu należy dodać $1.5 \cdot p_1 \vec{p}_2$. Jeśli samiec i samica znajdą się od siebie w odległości mniejszej niż $h_{p,♂,♀}/10$, to zaczynają się od siebie oddalać z prędkością $3 \cdot p_1 \vec{p}_2$ przez 5 jednostek czasu. Po 6 jednostkach czasu samica zaczyna składać $Poiss(6)$ jaj (każde co jedną jednostkę czasu, przez ten czas nie jest zainteresowana samcami, ani samce nie są zainteresowane nią), z których po 4 jednostkach czasu wylęgają się młode pardwy. Młode pardwy osiągają dojrzałość płciową po 15 jednostkach czasu.

3. Każda pardwa żyje w niewoli τ_p jednostek, gdzie τ_p jest zmienną losową z rozkładu wykładniczego z parametrem $\lambda = 70$ (czyli średnio 70 jednostek czasu).

4. Jeśli pardwa p_1 znajdzie się w odległości mniejszej niż $h_{p,w}$ od wilka w_1, to próbuje od niego uciec z szybkością $0.9 p_1 \vec{w}_1$, Jeśli znajdzie się w odległości mniejszej niż $h_{p,w}/10$, zostaje zjedzona.

Wilcy:

1. Wilcy trzymają się w stadzie. Jeśli odległość między dwoma dowolnymi wilkami (poza osobnikami alfa) jest większa niż h_{w_1,w_2}, to do dryfu należy dodać wektor $0.1 w_1 \vec{w}_2 \cdot (\rho(w_1,w_2) - 0.1)$.

2. Wśród wilków znajdują się osobniki alfa, które występują z prawdopodobieństwem 0.05. Inne wilki dążą do nich z prędkością $0.15 w_1 \vec{w}_2$, natomiast one

same są względem innych wilków (poza samicami i innymi osobnikami alfa) obojętne. Jeśli dwa samce alfa lub dwie samice alfa znajdą się od siebie w odległości mniejszej niż $h_{w,\alpha}$, do dryfu należy dodać wektor $-\vec{w_1 w_2}$.

3. Samce dążą do samic z prędkością $0.3\vec{w_1 w_2}$. Jeśli samiec i samica znajdą się od siebie w odległości mniejszej niż $h_{w,♂,♀}$, to zaczynają się od siebie oddalać z prędkością $2 \cdot \vec{w_1 w_2}$ przez 8 jednostek czasu. Po 10 jednostkach czasu samica rodzi $Poiss(2)$ młode co jedną jednostkę czasu, przez ten czas pozostaje obojętna dla samców. Młode osiągają dojrzałość płciową po 30 jednostkach czasu i trzymają się swojej matki z siłą $1.3 \cdot \vec{w_1 w_2}$.

4. Jeśli jakiś wilk złowi dorosłą pardwę, nakarmione nią zostają wszystkie wilki w promieniu h_{breed} od tego wilka oraz wszystkie jego niedojrzałe młode. Jeśli złowi młodą pardwę, karmi tylko siebie. Każdy wilk, który nie jadł nic od t_m jednostek czasu, umiera.

5. W niewoli wilki żyją τ_w jednostek czasu, gdzie $\tau_w = \mathcal{E}(200)$.

6. Wilcy mają czuły węch. Wabi je każda pardwa z siłą $\vec{w_1 p_1} \cdot (1 + \min\{0.3, 3/\rho^2(w_1, p_1)\})$.

Implementacja każdego z powyższych podpunktów daje 5 pkt. Dobierz parametry, które nie zostały wyżej ustalone, tak, aby populacja wilków i pardw w długim okresie czasu ($T = 70000$) nie wymierała ani nie wybuchała (20 pkt.). Wykonaj animację (nieskończoną, a nie .gif), w której pardwy i wilcy będą reprezentowane jako różnokolorowe kropki na płaszczyźnie (25 pkt.). Zadbaj o odpowiednią rozdzielczość czasową. Na animacji liczba pardw i wilków (płeć, młode osobniki, osobniki alfa) powinna być na bieżąco wyświetlana. Zadbaj o GUI: podczas wykonywania animacji powinna istnieć możliwość jej przeskalowania i przesuwania (3 pkt.). Animacja kończy się na polecenie użytkownika.

Opisz model Lotki-Volterry (5 pkt.). Czy nasze populacje pardw i wilków zachowują się zgodnie z jego przewidywaniami? Opisz krótko działanie i idee stojące za algorytmem *boids*, czyli tzw. algorytmem stada. Do czego się go używa? (3 pkt.) Postaraj się zoptymalizować program, np. wykonując obliczenia na karcie graficznej (5 pkt. za istotne usprawnienie).

3 Firma ubezpieczeniowa

Jesteś analitykiem w firmie ubezpieczeniowej sprzedającej ubezpieczenia OC i AC.

- Twój kapitał początkowy wynosi k_0.
- Zaczynasz z n sprzedanymi polisami OC, każda w cenie c_{OC} i $m = 0$, każda w cenie c_{AC}, sprzedanymi polisami AC.
- Na początku każdego miesiąca liczba kierowców, którzy rezygnują z OC, jest zmienną losową z rozkładu $B(N, (c_{OC}/m_{OC})^{2.5})$, gdzie N jest zmienną losową taką, że $P(N = n) = 0.5$, $P(N = floor(0.9n)) = 0.3$, $P(N = floor(n/2)) = 0.2$.

- Podobnie z kierowcami rezygnującymi z AC na początku każdego miesiąca: ich liczba jest zmienną losową z rozkładu $B(m, |Z| \cdot c_{AC}/m_{AC})$, gdzie Z jest zmienną losową z rozkładu $N(0,1)$, a m_{OC} i m_{AC} to liczby związane z popytem.

- Na początku każdego miesiąca zdobywasz nowych klientów. Liczba sprzedanych polis OC jest zmienną losową z rozkładu Poissona z parametrem $\mu \cdot \sqrt{M/c_{OC}} + s_{OC}/c_{OC}$, gdzie μ to nakłady na marketing (nie mogą przekroczyć całego posiadanego kapitału, o ich wielkości decyduje zarządca), a M ma następujący rozkład: $P(M = n/2) = 0.1$, $P(M = n/3) = 0.1$, $P(M = n/5) = 0.4$, $P(M = 50) = 0.4$, natomiast s_{OC} jest stałą związaną z popytem. Liczba sprzedanych polis AC z kolei jest zmienną z rozkładu $B(n, m/(n \cdot c_{AC}))$.

- Zgodnie z powyższym, na początku każdego miesiąca nasz kapitał wzrasta o $n \cdot c_{OC} + m \cdot c_{AC}$, ale musimy go pomniejszyć o μ.

- Szkody OC dla statystycznego kierowcy pojawiają się zgodnie z niejednorodnym procesem Poissona, a ich wielkość opisywana jest przez rozkład Pareto. Funkcja intensywności $\lambda_0(t)$ w procesie Poissona powinna zostać opisana na podstawie rzeczywistych danych (jaki jest rozkład intensywności wypadków samochodowych w określonych miesiącach, dniach tygodnia? jak dni świąteczne wpływają na liczbę wypadków samochodowych? Przytoczyć odpowiednie dane). Wartość referencyjna λ_0 w jakiś „typowy” dzień powinna wynosić 1 i zmieniać się w zależności od mniej lub bardziej ryzykownych okresów w ciągu roku.

- λ_0 nie jest jednak intensywnością sygnałów o szkodach, jakie pojawiają się w naszej firmie. Szkody będą pojawiać się tym częściej, im więcej sprzedamy polis. Ostateczna funkcja intensywności będzie swego rodzaju podwójnie stochastycznym procesem Poissona:

$$\lambda(t) = \sum_{i=0}^{n(t)} W_i(t)\lambda_0(t), \tag{1}$$

  gdzie $n(t)$ jest liczbą sprzedanych w danym miesiącu polis, a W_i jest zmienną losową opisującą, jak często dany kierowca miewa stłuczki (zakładamy, że W_i z rozkładu $E(0.05)$ są niezależne; w każdym miesiącu będziemy je generować od nowa)

- Szkody pojawiające się w związku ze sprzedażą pojedynczej polisy AC są procesem odnowy, w którym czas oczekiwania na sygnał T_{AC}^i są zmiennymi z rozkładu $U(5, 5+V_i)$, gdzie V_i są niezależne i pochodzą z rozkładu Rayleigha o średniej 100. Wielkość szkód opisywana jest przez rozkład $\min\{|C_i|/2, 500 \cdot c_{AC}\}$, gdzie C_i są niezależne ze standardowego rozkładu Cauchy'ego. $10000 \cdot c_{AC}$ oznacza tyle, że w umowie zastrzegamy sobie, że wielkość wypłaty nie może przekroczyć pewnej wielokrotności ceny polisy.

- W przypadku, gdy nasz kapitał spadnie poniżej zera, możemy zaciągnąć kredyt w wysokości nie większej niż $120n$. Zaczynamy go spłacać pół roku od zaciągnięcia, 20 rat miesięcznych po $1/12$ wysokości zaciągniętego kredytu każda.

Kredyt możemy wziąć tylko raz w roku i tylko w wypadku, gdy nie mamy żadnych zobowiązań.

- Jeśli wartość naszego kapitału spadnie poniżej $-2k_0 - nc_{AC}$, automatycznie Bankrutujemy.

- Ruina paryska. Bankrutujemy także, jeśli nasz kapitał pozostaje stale ujemny przez dwa lata.

Napisz funkcję symulującą proces ryzyka opisany powyższymi warunkami (za każdy po 5 pkt.). Podobnie jak w poprzednim zadaniu, napisz nieskończoną animację z sensownym GUI (15 pkt). Dobierz sensownie parametry niezależne od strategii (s_A, m_A, λ, k_0, parametry używanych rozkładów) (15 pkt.). Co to są rozkłady a priori i a posteriori? (3 pkt.) Zbadaj rozkłady a posteriori liczby sprzedanych polis. (7 pkt.) Znajdź przykłady sytuacji, w których rozkład Pareto dobrze modeluje jakieś zjawisko (podaj referencje). (3 pkt.) Napisz funkcję do generowania zmiennych z rozkładu Pareto. (3 pkt.)

Dyrektor firmy ubezpieczeniowej zlecił Ci zaproponować odpowiednie strategie na marketing, zaciąganie kredytów i ustalanie wysokości cen. Zaproponuj w miarę sensowne strategie (20 pkt.) i przeprowadź dla dyrektora „analizy" (dyrektor nie sprecyzował, o co mu chodzi, ale możesz podejrzewać, że na pewno będzie zainteresowany prawdopodobieństwem bankructwa w ciągu najbliższych 10 lat, oczekiwanym rocznym zyskiem, wpływem strategii na zysk i ryzyko, kiedy nasz dyrektor zostanie milionerem itd.) (30 pkt.).

4 Czasy trafienia w zbiór

Co to są momenty stopu (momenty Markowa)? (5 pkt.) Co to znaczy, że proces stochastyczny trafia w punkt? Co to są procesy tranzytywne, zbiory polarne? (2 pkt)

1. Oszacuj średni czas wyjścia d-wymiarowego ruchu Browna ze zbiorów: $B_d(0,1)$, $[-1,1]^d$ (w szczególności: jak zmienia się średni czas wyjścia w zależności od d?) Zbadaj rozkład tych czasów wyjścia dla $d = 1,2,3$.(7 pkt.)

2. Zbadaj prawdopodobieństwo, że do chwili T ruch Browna w $\mathbb{R}^2$ po wyjściu z $B((0,0),5)$ wróci do $B(0,0),0.01)$. (5 pkt.)

3. Oszacuj prawdopodobieństwo trafienia ruchu Browna w $\mathbb{R}^2$ w zbiór $B((5,5),1)$ do chwili $T = 100000$. (4 pkt.)

4. Oszacuj prawdopodobieństwo, że ruch Browna w $\mathbb{R}^2$ trafi w $B((2,2),1)$, nie trafi w $B((3,3),0.5)$ i następnie trafi w $B((-1,3,1.6)$ do chwili $T = 1000000$. (4 pkt.)

5 Zadanie do kupienia

Poniższe zadania wiążą się z tematyką symulowania procesów stochastycznych. Część z nich poznacie lepiej na innych kursach w przyszłości. W niniejszym sprawozdaniu chodzi o wstęp do wymienionych tematów. W większości przypadków należy napisać krótki referat oraz przeprowadzić odpowiednie symulacje, do których trudniejsze partie kodu podam.

5.1 Zasady licytacji

Spośród podanych niżej podpunktów każda grupa dostanie tylko jeden. Spośród grup, które zgłoszą się do danego podpunktu, wybrana zostanie ta, która kupi je za największą liczbę plusów. Opłatę plusową pobiera się od tej grupy, która zwycięży w danej licytacji, przy czym to, której osobie zostaną odjęte plusy, ustala grupa. W przypadku takich samych deklaracji, grupa zostanie wybrana losowo. Przykład: do podpunktu zgłosiły się grupy A,B,C,D,E. Grupy A i B nie chcą płacić plusami za ten podpunkt. Grupa C deklaruje 3 plusy: 1 od jednej osoby i 2 od innej. D i E deklarują 4 plusy. Zadanie rozlosowywane jest między D i E. Wygrywa E, w której są 3 osoby. Ponieważ plusy zdobywane były nierównomiernie (otrzymują je tylko osoby obecne na zajęciach), grupa E postanowiła, że jedna osoba odda 3 swoje plusy, jedna 1 plusa, a jedna żadnego. Pozostałe grupy nie płacą za ten podpunkt nic.

Każda grupa może się zgłosić do co najwyżej pięciu podpunktów. Jeżeli przegra we wszystkich licytacjach, otrzyma jeden z podpunktów, do którego się nie zgłosiła.

Licytacja kończy się 12.05. Ze względu na to, że każda z grup ma dostęp do Facebooka, aktualne ceny zadań będą przedstawiane na grupie. Treści zadań będą się pojawiały sukcesywnie do końca maja, na razie muszą wystarczyć ogólne tematy: podobnie jak w przypadku pracy inżynierskiej lub magisterskiej, dokładny zakres pracy zostaje ustalony w podczas konsultacji z promotorem.

5.2 Procesy α-stabilne

5.3 Model ARMA

5.4 Łańcuchy Markowa: patrolling

5.5 Łańcuchy Markowa: czasy pokrycia

5.6 Wstęp do symulowania sieci społecznościowych

5.7 Spektrum procesu

5.8 Branching processes: modele rozwoju kolonii bakteryjnych

5.9 Arytmetyka zmiennoprzecinkowa, czyli na co uważać przy symulacjach

5.10 Czasy trafienia w zbiór wielowymiarowego procesu Wienera o niejednostkowej macierzy kowariancji i z dryfem

5.11 Statystyka procesów stochastycznych. Podstawowe narzędzia.

5.12 Autobusy i pasażerowie. Niejednorodny proces Poissona w praktyce.

5.13 Analiza złożoności algorytmów. Jak optymalizować programy w języku Python/Matlab?

5.14 Julia - język do obliczeń naukowych

5.15 Równanie Langevina. Równanie Fokkera-Plancka. Związki równań różniczkowych z probabilistyką.

5.16 Implementacja wybranych testów statystycznych.

5.17 Małpa, którą zwą Szekspir

5.18 Awaria elektrowni atomowej

5.19 Obliczenia równoległe a zakleszczenia. Jak bezpiecznie optymalizować symulacje?

5.20 Procesy z długą pamięcią.

5.21 Jak rosną drzewa? Symulacje procesów stochastycznych jako narzędzie do rysowania krzewów.