Вериги на Марков

Регулярни вериги на Марков

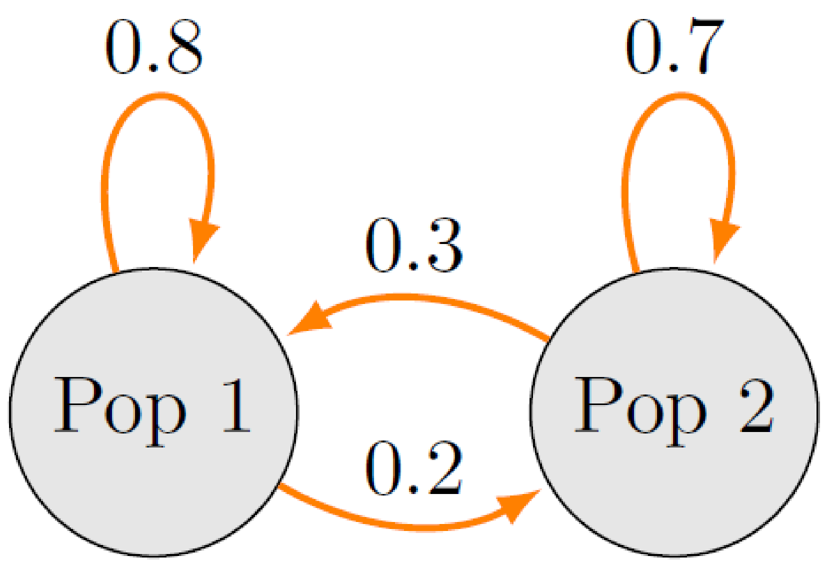
- *Веригите на Марков* носят името на руския математик *Андрей Андреевич Марков* (1856-1922).
- Марков изучавал фонетичната структура на руския език и търсил статистически методи за анализ на характеристичните особености на стиловете на различни писатели.
- За изследванията си Марков използвал откъс от 20 000 последователни букви от текста на романа "Евгений Онегин" на Ал. Пушкин, като анализирал броя на гласните (Г) и съгласните (С) в текста и тяхната последователност - вероятността за преходите между двата вида звуци: Г→Г, Г→С, С→Г, С→С.
- Марков установил, че преходите Г→С, С→Г се срещат почесто от очакваното, изчислено чрез процентното съдържание на гласни и съгласни в текста (което не е учудващо, имайки предвид, че езикът е създаден за говорене, а съединяването на съгласни с гласни звуци улеснява изговарянето).
- www.americanscientist.org/article/first-links-in-the-markov-chain

Верига на Марков е вероятностен (стохастичен) процес, описващ система, която във всеки момент от време се намира в едно от краен брой състояния. На всяка стъпка системата преминава от едно състояние в друго, като преходът между състоянията се осъществява на равни дискретни интервали от време, по случаен начин и вероятността за преход не зависи от времето, а само от текущото състояние.

Горното определение е за един вид вериги на Марков - *крайни дискретни хомогенни вериги на Марков*. Броят на състоянията, в които може да се намира системата, е краен, а преходите между състоянията се осъществяват на равни дискретни интервали от време, като вероятностите за преход не се изменят с течение на времето.

Пример 1. Разглеждаме миграционен модел между две популации на един и същи вид. Известно е, че всяка година 80% от индивидите в първата популация остават в нея, а останалите 20% мигрират към втората популация и 70% от индивидите от втората популация остават в нея, а 30% мигрират към първата.

Графично моделът може да се представи чрез следната преходна диаграма (претеглен ориентиран граф - вероятността за всеки преход е записана като "тегло" над съответното ориентираното ребро).



Нека в началното състояние в двете популации имаме по 100 индивида. Векторът на началното състояние е $v_0 = (100, 100)$.

При дадените вероятности за преход между популациите, в края първата година, в първата популация ще са останали 80 от първоначалните индивиди и към тях ще са се присъединили 30 от втората популация. Във втората популация ще са останали 70 от първоначалните индивиди и към тях ще се добавят 20 от първата популация. По този начин в края на първата година броят на индивидите в двете популации ще бъде съответно 110 и 90 (в модела предполагаме, че не се добавят нови индивиди и не измират). Нека означим $v_1 = (110, 90)$.

Чрез матричния апарат моделът може да бъде представен по следния начин. Нека формираме квадратната матрица от 2-ри ред P, в която запишем (по редове) вероятностите за преход между двете популации (състояния) в модела

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}.$$

Матрицата $P = (p_{ij})$ се нарича ***преходна матрица*** на веригата на Марков.
Сумата от елементите във всеки ред на P е равна на 1, тъй като това е сумата от вероятностите на всички възможни събития, които могат да се случат, ако процесът се намира в състоянието с пореден номер, равен на номера на съответния ред.

Квадратна матрица, всички елементи на която имат стойности между 0 и 1, като сумата на елементите от всеки ред (стълб) е равна на 1, се нарича *стохастична (вероятностна) матрица.*

Тогава векторът v_1, отговарящ на числеността на двете популации след първата година, се получава чрез

$$v_1 = v_0 P = \begin{pmatrix} 100 & 100 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 110 & 90 \end{pmatrix}.$$

Аналогично, числеността на двете популации в края на втората година (при условие, че дадените вероятности за преход се запазят) може да се получи чрез

$$\begin{aligned} v_2 &= v_1 P = (v_0 P) P = v_0 P^2 = \\ &= \begin{pmatrix} 100 & 100 \end{pmatrix} \begin{pmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{pmatrix} = \begin{pmatrix} 115 & 85 \end{pmatrix}. \end{aligned}$$

След k на брой години, при същите вероятности за преход, за броя на индивидите в двете популации ще имаме

$$v_k = v_{k-1} P = v_0 P^k.$$

Елементите на P^2 ни дават информация за вероятността процесът да се намира в дадено състояние след 2 последователни прехода. По-точно, елементът в i-тия ред и j-тия стълб на P^2 е равен на вероятността процесът след 2 прехода да се намира в състояние j, ако е започнал от състояние i. Аналогично важи за елементите на P^k, $k \geq 1$.

Важен въпрос е - може ли да се направи дългосрочна прогноза за поведение на процеса?
Ще разгледаме няколко вида крайни дискретни вериги на Марков, за които могат да бъда направени такива прогнози.

Онлайн приложение за конструиране и изследване на крайни дискретни марковски вериги с 2 състояния:
https://setosa.io/ev/markov-chains/.

Вериги на Марков в тази тема:

- *ергодична верига* - за всеки елемент p_{ij} на преходната матрица P съществува естествено число k такова, че елементът на същата позиция в P^k е по-голям от 0, т.е. от всяко състояние може да се отиде във всяко друго (не непременно чрез един преход). Еквивалентно условие - за съответния на матрицата P ориентиран граф е изпълнено условието, че за всяка наредена двойка (i, j) от върхове съществува ориентиран път от i към j (нарича се *силно свързан ориентиран граф*, а матрицата P се нарича *неразложима*).

- *регулярна верига* - съществува $k \in \mathbb{N}$ такова, че всички елементи на P^k са по-големи от 0. Примерът за двете популации, който разгледахме, е на такава верига на Марков. Всяка регулярна верига е ергодична. Матрица P с елементи неотрицателни числа, за която съществува k такова, че всички елементи в P^k са положителни числа, се нарича *примитивна* (*апериодична*).

Какво означава ориентиран граф, съответен на квадратна матрица с елементи неотрицателни числа (нарича се *неотрицателна матрица*) и в частност на стохастична матрица?

Нека $A = (a_{ij})$ е квадратна матрица от n-ти ред и $a_{ij} \geq 0$. Разглеждаме ориентиран граф с n на брой върха, номерирани от 1 до n, като

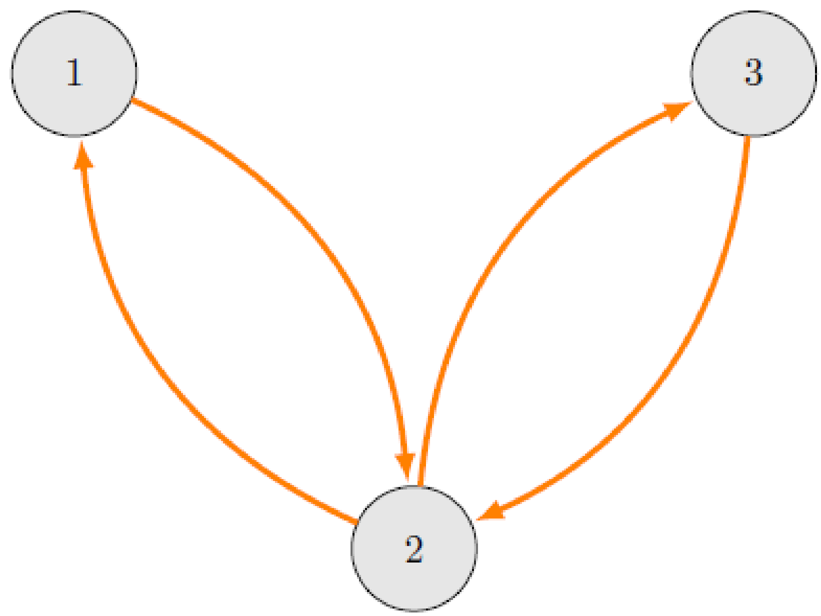
- ако $a_{ij} > 0$, то съществува ориентирано ребро от върха с пореден номер i към върха с пореден номер j;
- ако $a_{ij} = 0$, то не съществува ориентирано ребро от върха с пореден номер i към върха с пореден номер j.

Ориентираният граф, дефиниран съгласно горното правило, се нарича съответен на неотрицателната матрица A.

Пример 2. Нека е дадена неотрицателната квадратна матрица

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}.$$

Ориентираният граф, съответстващ на P, има 3 върха и ориентирани ребра съществуват между наредените двойки върхове $(1, 2)$, $(2, 1)$, $(2, 3)$ и $(3, 2)$.



Тъй като сумата на елементите от всеки ред на неотрицателната матрица P е равна на 1, то P е стохастична матрица (матрица на верига на Марков).
От диаграмата се вижда, че ориентираният граф, съответен на P, е силно свързан. Следователно веригата на Марков с преходна матрица P е ергодична (P е неразложима матрица).

Можете да проверите, че

$$P^2 = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \end{pmatrix}, \quad P^3 = P, \quad P^4 = P^2,$$

откъдето $P^{k+2} = P^k$ за всяко $k \geq 1$, т.е. се наблюдава периодичност в степените на P. Следователно не съществува степен на P, за която всички елементи да са положителни числа. Следователно марковската верига с преходна матрица P не е регулярна.
Най-малкото естествено число h, за което $P^{k+h} = P^k$ за всяко $k \geq 1$, се нарича *период* на P.

Векторът $v = (x_1, x_2, ..., x_n)$ се нарича ***вероятностен (стохастичен) вектор***, ако

$$0 \leq x_i \leq 1, \qquad x_1 + x_2 + ... + x_n = 1.$$

Примери: $(0.5, 0.5)$, $(1, 0, 0)$, $(0.25, 0.5, 0.25)$. Чрез вероятностни вектори се описват състоянията на верига на Марков.

Нека се върнем на примера с регулярната верига на Марков, описващата миграционен модел между две популации с преходна матрица

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}.$$

Нека стартираме процеса с вектор на началното състояние $v_0 = (1, 0)$, т.е. процесът се намира в състояние № 1.

Интересуваме се как ще се развива процесът в бъдеще (след голям брой последователни преходи между състоянията), дали може да се направи прогноза, дали има устойчива тенденция.

$$v_1 = v_0 P = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \end{pmatrix}$$

$$v_2 = v_1 P = v_0 P^2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}$$

$$v_3 = v_2 P = v_0 P^3 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.65 & 0.35 \\ 0.525 & 0.475 \end{pmatrix} = \begin{pmatrix} 0.65 & 0.35 \end{pmatrix}$$

$$v_{10} = v_0 P^{10} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.600391 & 0.399609 \\ 0.599414 & 0.400586 \end{pmatrix} = \begin{pmatrix} 0.600391 & 0.399609 \end{pmatrix}$$

$$v_{21} = v_{20} P = v_0 P^{21} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.4 \end{pmatrix}$$

$$v_{100} = v_{99} P = v_0 P^{100} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.4 \end{pmatrix}.$$

Виждаме, че започвайки от първото състояние, след даден момент нататък вероятностите процесът да бъде във всяко от състоянията не се променят (стават постоянни).

Нещо повече. Оказва се, че същото ще се случи независимо от началното състояние v_0. За всеки начален вероятностен вектор v_0 съществува m такова, че при $k \geq m$ имаме $v_k = v_{k+1} = \ldots$.

Съществува вероятностен вектор $\widehat{v}$ такъв, че

$$\lim_{k\to\infty} v_k = \lim_{k\to\infty}(v_0 P^k) = v_0(\lim_{k\to\infty} P^k) = \widehat{v}.$$

Векторът $\widehat{v}$ се нарича ***гранично разпределение*** на марковската верига. Не всяка марковска верига има гранично разпределение. Регулярните вериги имат, ергодичните периодични вериги - не (като тази от Пример 2). Регулярните вериги се наричат още апериодични ергодични.

От експеримента ни забелязахме, че за разглежданата от нас регулярна марковска верига съществува m такова, че при $k \geq m$ имаме $v_k = v_{k+1} = v_{k+2} = \dots$, т.е. съществува вектор π, за който

$$\pi = \pi P.$$

Векторът π се нарича ***стационарно разпределение (стационарен вектор)*** на веригата на Марков. Очевидно той е ляв собствен вектор на матрицата P, отговарящ на собствената стойност $\lambda = 1$.

$\lambda = 1$ е собствена стойност на всяка стохастична матрица. Изпълнено е $|\lambda| \leq 1$ за всяка собствена стойност λ на произволна стохастична матрица.

За всяка ергодична верига на Марков (следователно и за всяка регулярна) съществува единствен вероятностен собствен вектор, отговарящ на $\lambda = 1$. Това твърдение е следствие от Теоремата на Перон-Фробенуис.

За регулярните вериги на Марков стационарното разпределение съвпада с граничното, съвпада и с всеки от редовете на матрицата

$$M = \lim_{k\to\infty} P^k.$$

За веригата от нашия пример стационарното разпределение е $\pi = (0.6, 0.4)$. Нека се убедим в това, като намерим собствените стойности на преходната матрица P. Характеристичното ѝ уравнение е

$$\begin{vmatrix} \frac{4}{5}-\lambda & \frac{1}{5} \\ \frac{3}{10} & \frac{7}{10}-\lambda \end{vmatrix} = 0 \quad \Leftrightarrow \quad (4-5\lambda)(7-10\lambda)-3=0,$$

еквивалентно на $2\lambda^2-3\lambda+1=0$, чиито корени са $\lambda_1 = 1$, $\lambda_2 = \frac{1}{2}$. Нека $u_1 = (x, y)$ е ляв собствен вектор, отговарящ на $\lambda_1 = 1$. Тогава

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -\frac{1}{5} & \frac{1}{5} \\ \frac{3}{10} & -\frac{3}{10} \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \quad \Leftrightarrow \quad 2x-3y=0,$$

чието решение са векторите от вида $(3p, 2p)$. При $p = 1$ имаме $u_1 = (3, 2)$.

За да получим вероятностен вектор от вектора $u_1 = (3, 2)$, трябва да го нормираме по т.нар. единична норма (манхатънска норма), която за произволен вектор $v = (x_1, x_2, ..., x_n)$ се дефинира чрез

$$||v||_1 = |x_1| + |x_2| + ... + |x_n|.$$

Тогава векторът $\frac{u}{||u||_1}$ е вероятностен (при условие, че $0 \leq x_i \leq 1$). Следователно трябва да разделим координатите на u_1 на сумата на координатите му. Така получаваме вектора
$\pi = \left(\frac{3}{5}, \frac{2}{5}\right) = (0.6, 0.4)$.

С този вектор се описва дългосрочното поведение на процеса, т.е. 60% от времето процесът ще се намира в състояние № 1 и 40% в състояние № 2. Или казано по друг начин в контекста на примера - 60% от общия брой индивиди на двете популации ще се намират в първата популация и 40% във втората.

Теорията на стохастичните матрици е едно от приложенията на по-широка теоретична рамка, известна като теория на неотрицателните матрици. Централна роля в тази теория играе Теоремата на Перон-Фробениус.

Теорема (Перон-Фробениус) *Нека $A = (a_{ij})$ е квадратна неотрицателна матрица, т.е. $a_{ij} \geq 0$, която е неразложима. Тогава:*

- *A има единствена собствена стойност $\lambda > 0$, за която $|\lambda_i| \leq \lambda$ за всяка друга собствена стойност λ_i на A;*
- *λ е прост корен на характеристичния полином на A, т.е.* $\mathrm{alg\ mult}(\lambda) = \mathrm{geom\ mult}(\lambda) = 1$*;*
- *На λ съответства единствен собствен вектор v с координати положителни числа и такъв, че $||v||_1 = 1$ (нарича се вектор на Перон, важи и за десен, и за ляв собствен вектор);*
- *Ако A е примитивна неразложима матрица, то $|\lambda_i| < \lambda$ за всяка друга собствена стойност λ_i на A.*

Теоремата е доказана първоначално за матрици с елементи положителни числа (*положителни матрици*) през 1907 г. от немския математик Оскар Перон и обобщена през 1912 г. за неотрицателни неразложими матрици от Фердинанд Георг Фробениус.

От гледна точка на веригите на Марков теоремата е в сила за ергодични вериги, като $\lambda = 1$. Последното условие от теоремата е изпълнено само за регулярни вериги на Марков.

Ако се върнем съм марковската верига от Пример 2, която е ергодична, но не е регулярна, ще забележим, че тя няма гранично разпределение, тъй като редицата от степените на преходната матрица не е сходяща. Тъй като $P^{k+2} = P^k$, $k \geq 1$, то тази редица има вида

$$P, \quad P^2, \quad P, \quad P^2, \ldots$$

Стационарното разпределение е единствено и е вероятностният ляв собствен вектор, съответстващ на собствената стойност 1. Намерете собствените стойности на преходната матрица

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}.$$

и установете, че те са $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 0$.

Нека $v_1 = (x, y, z)$ е ляв собствен вектор, съответстващ на $\lambda_1 = 1$. Следователно

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix},$$

еквивалентно на системата

$$\left| \begin{array}{l} -2x + y = 0 \\ x - y + z = 0 \\ y - 2z = 0. \end{array} \right.$$

Решенията на горната система са вектори от вида $(p, 2p, p)$, откъдето $v_1 = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$ е стационарното разпределение на тази марковска верига.

Литература

1. Д. Мекеров, Н. Начев, Ст. Миховски, Е. Павлов, *Линейна алгебра и аналитична геометрия*, Пловдив, 1997.

2. D. C. Lay, S. R. Lay, Judi J. McDonald, *Linear algebra and its applications*, 5th ed. Pearson, 2016.

3. G. Strang, *Linear algebra and its applications*, 4th ed., Nelson Engineering, 2007, ISBN-13: 978-813-150-172-6.

4. H. Anton, C. Rorres, *Elementary Linear Angebra (applications version)*, 11th ed., Wiley, 2014, ISBN 978-1-118-43441-3.

5. S. Axler, *Linear Algebra Done Right*, 3rd ed., Springer, 2015.

6. K. Singh, *Linear Algebra Step by Step*, Oxford University Press, 2014.

7. C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

8. S. J. Leon, *Linear Algebra with Applications*, 9th ed., Pearson, 2015.