Markov Lánc - Monte Carlo módszerek

Valószínűségszámítási módszerek

2021. április 15.

Eötvös Lóránd Tudományegyetem - Természettudományi Kar

Egyszerű példa a Markov láncokra

Budapesti közlekedési modell

Tegyük fel, hogy Budapesten az emberek vagy tömegközlekednek vagy autóznak. A kezdeti évben az emberek 20 százaléka autózik és 80 százaléka tömegközlekedik. Minden következő évben az autósok 5 százaléka vált tömegközlekedésre és a tömegközlekedők 10 százaléka vált autóra. Milyen lesz az arányuk hosszútávon?

Megoldás mátrixokkal

Algebra II.-ben tanultuk.

Legyen x az autósok aránya, y a tömegközlekedők aránya. A n. évi adatot megkaphatjuk az alábbi mátrixszorzás segítségével:

$$\begin{bmatrix} 0.95 & 0.1 \\ 0.05 & 0.9 \end{bmatrix}^n \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

Ezt hatványozva a következő egyenletet kapjuk

$$\begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

mivel $x + y = 1$ ezért ezek tetszőleges választása mellett erre az eredményre jutunk. Az egy fixponttal rendelkező Markov Láncokat stacionáriusnak is nevezzük.

Def.: Egy ML akkor stacionárius, ha $P(x, y) = (x, y)$.

Első rendű Markov láncok definíciója

Csak a jelenlegi állapottól függ a következő lépésbeli érték, a korábbi értékek nem befolyásolják. Ilyen például az érmedobás vagy akár az előbbi közlekedési modell.

Definíció

Véges (vagy megszámlálhatóan végtelen) sok értéket felvevő $X_1, ..., X_t ...$ valószínűségi változók sorozata akkor alkot Markov-láncot, ha bármely t kezdőpillanat esetén a sorozat t utáni viselkedése csak a t időpontbeli X_t értéktől függhet.

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = P(X_{n+1} = j | X_n = i_n)$$

d-ad rendű esetben csak az utolsó d állapottól függ az érték.

Példák a Markov láncokra

1. Bolyongás a Lágymányoson
2. Google PageRank
3. Populációmodellek
4. Véletlen bolyongás (folytonos állapoterű)
5. Véletlen mintavételezés (MCMC)

Mire figyeljünk a modellben?

- Rajzoljunk fel egy irányított gráfot és utána annak a szomszédsági mátrixát
- Az a_{ij} elem azt jelenti, hogy mekkora a valószínűsége, hogy a következő körben a j. csúcsból az i. csúcsba megyünk-
- Minden él nemnegatív és minden csúcsra a kimenő élek összege 1 (ezt a mátrix oszlopai jelentik)
- A diagonálelemek az egy helyben maradás valószínűségét jelölik

Nem csak véges állapottéren lehet értelmezni, tetszőleges mérhető térben működik. Milyen problémák lehetnek? Ha vannak olyan pontok, amikből nem lehet eljutni másikba. Pozitív, amennyiben minden pontból minden másik elérhető pozitív valószínűséggel. k-periodikus, amennyiben egy állapotba csak a k valamilyen többszörövésvel lehet visszajutni és aperiodikus, ha ezen periódusok legnagyobb közös osztója 1. Egy Markov lánc akkor ergodikus, ha aperiodikus és pozitív.

Monte Carlo módszerek

Sokszor analitikusan nehezen megoldható egy feladat. Egy százdimenziós térfogatot kiszámolni például reménytelen a legtöbb esetben. De akkor is "segíthet", ha egyszerűbb esetekben nem ismernénk a megoldást. Független azonos eloszlású változókat használva azonban jól közelíthető a várható érték. Ha jól generálunk véletleneket, akkor tudjuk őket használni a modellben.

Egyszerű események várható értékének becslése véletlennel

Minek nagyobb az esélye, hogy három kockával egyszerre dobva a számok összege 9 vagy 10?

Esetszétválasztással megkaphatjuk, hogy a valószínűségek 25/216 illetve 27/216. De ha nem szeretünk számolni? Próbáljuk meg ezt a Matlab kódot.

Egyszerű események várható értékének becslése véletlennel

Minek nagyobb az esélye, hogy három kockával egyszerre dobva a számok összege 9 vagy 10?
Esetszétválasztással megkaphatjuk, hogy a valószínűségek 25/216 illetve 27/216. De ha nem szeretünk számolni? Próbáljuk meg ezt a Matlab kódot.

Az n-gömb térfogatának kiszámítása

A kockázásnál a szimulációk száma meghaladta az összes esetet
Monte Carlo módszer ereje, hogy kisebb minta esetén is közelíti az eloszlás várható értékét.
Nézzük az n dimenziós egységgömb térfogatát különböző esetekre.

Térfogatok

Dim	Monte Carlo	Konkrét	Hiba (%)
2	3,1408	3,1416	0,02546
3	4,1882	4,1888	0,01432
5	5,2644	5,2638	0,0114
10	2,5815	2,5502	1,22735
15	0,3277	0,3814	14,07971
18	0,0524	0,0821	36,17539

A pontosság növelhető az interációk számával, de egy idő után már itt is reménytelen.
Idáig független mintákat vettünk az eloszlásokból. Mi történik, ha az X_{i+1} minta függ X_i-től?

Markov Chain Monte Carlo

Tegyük fel, hogy van egy valószínűségi változónk, amiből szeretnénk mintavételezni. Ezt a változót nem ismerjük, csak egy konstans szorzó erejéig. Ekkor az alábbiakat tehetjük.

- Vegyünk egy tetszőleges kiinduló x elemet az eseménytérből, minden elemhez $f(x)$ értéket rendeljük hozzá.
- Vegyünk valamilyen g eloszlást x körül, amiből könnyen tudunk mintavételezni. Az így kapott y elem lesz a következő javasolt állapot.
- Ezt a javaslatot a következő valószínűséggel fogadjuk el

$$\alpha(x, y) = \min\left(1, \frac{f(y)g(x|y)}{f(x)g(y|x)}\right)$$

$$\alpha(x, y) = \min\left(1, \frac{f(y)g(x|y)}{f(x)g(y|x)}\right)$$

Ha a tört kifejezése nagyobb, mint egy, akkor biztosan elfogadjuk y-t. Amennyiben kisebb, mint 1, akkor $\alpha(x, y)$ valószínűséggel fogadjuk el y-t. Amennyiben nem fogadjuk el, akkor az új állapot meg fog egyezni a korábbival.

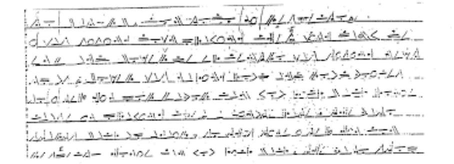
A g függvény, amennyiben szimmetrikus, akkor az egyenlet egyszerűsíthető $\alpha(x, y) = min\left(1, \frac{f(y)}{f(x)}\right)$

Ez az algoritmus egy Markov láncot generál, ami konvergál az f stacionárius eloszlásához. Az első néhány iterációt általában elhagyjuk, mivel pontatlanok (hiszen tetszőleges kiinduló értéket választottunk)

(https://youtu.be/U561HGMWjcw?t=469)

Példa az alkalmazásra

Egy helyettesítő rejtjelezést szeretnénk visszafejteni. Minden szimbólum egy karaktert helyettesít. Az f függvény a kódolt szöveg minden eleméhez meghatároz egy karaktert. Találjuk meg azt az f függvényt, amivel a szöveg értelmes.



Folytatás

Ahhoz, hogy össze lehessen hasonlítani két f függvényt, vezessünk be egy $Pl(f)$ értéket. Jelentse $M(x, y)$ az egymást követő (x, y) szimbólumok arányát. Ehhez vegyünk egy referencia szöveget, mondjuk a Háború és békét és legyen $Pl(f) = \prod_i M(f(s_i), f(s_{i+1}))$ vagyis a kódolt szöveg s_i, s_{i+1} karaktereire nézzük meg, hogy mekkora valószínűséggel szerepelnek egymást mellett a referencia szövegben.

- Induljunk ki egy tetszőleges f függvényből
- Számoljuk ki rá a $Pl(f)$-et
- Válasszunk ki a kódolt szövegből két szimbólumot és cseréljük fel a hozzájuk rendelt értékeket. Legyen ez f^*
- Amennyiben $Pl(f^*) > Pl(f)$ elfogadjuk, amennyiben f^* kisebb $\frac{Pl(f^*)}{Pl(f)}$ valószínűséggel fogadjuk el.

Így biztosak lehetünk benne, hogy nem ragadunk egy pontban. Kb 10000 iteráció szükséges a használható eredményhez.

Visszafejtett szöveg

to bat-rb. con todo mi respeto. i was sitting down playing chess with danny de emf and boxer de el centro was sitting next to us. boxer was making loud and loud voices so i tell him por favor can you kick back homie cause im playing chess a minute later the vato starts back up again so this time i tell him con respecto homie can you kick back. the vato stop for a minute and he starts up again so i tell him check this out shut the f**k up cause im tired of your voice and if you got a problem with it we can go to celda and handle it. i really felt disrespected thats why i told him. anyways after i tell him that the next thing I know that vato slashes me and leaves. dy the time i figure im hit i try to get away but the c.o. is walking in my direction and he gets me right dy a celda. so i go to the hole. when im in the hole my home boys hit doxer so now "b" is also in the hole. while im in the hole im getting schoold wrong and

Források

- Kelemen Kitti: Az MCMC algoritmus és néhány alkalmazása (2017)
- Persi Diaconis: The Markov Chain Monte Carlo Revolution
- Ben Lambert: An introduction to the Random Walk Metropolis algorithm (https://youtu.be/U561HGMWjcw)
- ritvikmath: Markov Chain Monte Carlo (MCMC): Data Science Concepts (https://youtu.be/yApmR-c_hKU)
- ritvikmath: Metropolis - Hastings: Data Science Concepts (https://youtu.be/yCv2N7wGDCw)

Köszönöm a figyelmet!