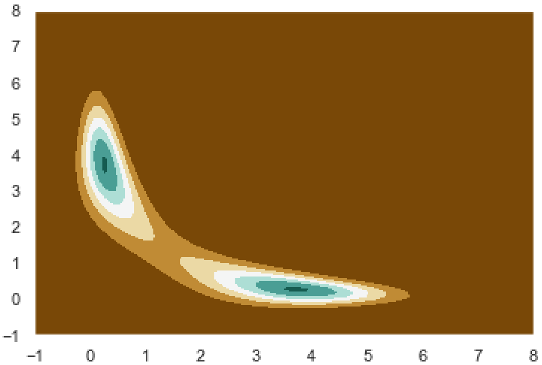
Próbnik Gibbsa

Natalia Safiejko

Plan prezentacji

Co to?

Algorytm próbnika Gibbsa jest techniką numeryczną stosowaną do symulowania złożonych rozkładów prawdopodobieństwa. Algorytm ten polega na wykorzystaniu warunkowego rozkładu prawdopodobieństwa, aby wygenerować próbkę z pełnego rozkładu. Działanie algorytmu opiera się na wykonaniu sekwencji kroków, w których każdy krok wykorzystuje jedną zmienną, aby wygenerować nową wartość próbki.

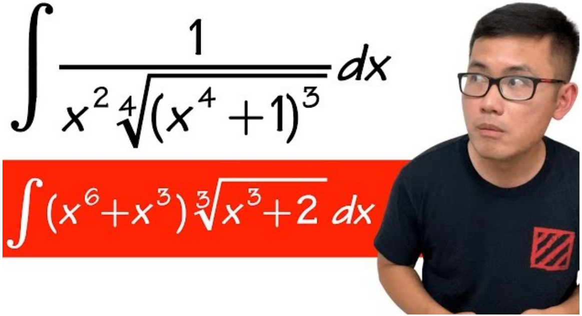


Zastosowania

Algorytm próbnika Gibbsa jest wykorzystywany w wielu dziedzinach, takich jak analiza danych, sztuczna inteligencja, bioinformatyka, chemia i fizyka. Jego zastosowania obejmują symulacje molekularne, estymację parametrów modeli statystycznych, analizę sieci neuronowych, detekcję zmian w sygnałach, klasyfikację i rozpoznawanie obrazów oraz generowanie danych losowych.

Kiedy używamy?

Próbnik Gibbsa jest używany w sytuacjach, gdy nie jest możliwe uzyskanie analitycznej formuły rozkładu prawdopodobieństwa, a także wtedy, gdy rozkład ten jest skomplikowany i nie ma wystarczająco dużo zasobów obliczeniowych, aby go dokładnie zasymulować. W takich sytuacjach próbnik Gibbsa umożliwia generowanie próbek z rozkładu prawdopodobieństwa poprzez symulację warunkową, co pozwala na obliczenie różnych parametrów i statystyk.



NA PRZYKŁAD

Mamy dany wielowymiarowy rozkład prawdopodobieństwa. Zależy nam na otrzymaniu np. średniej czy wartości oczekiwanej z gęstości brzegowej.

$$f(x) = \int \cdots \int f(x, y_1, \ldots, y_p) dy_1 \ldots dy_p$$

Przykład na dwuwymiarowej zmiennej

Generowanie próbek

X,Y - zmienne losowe

Generowanie próbek z rozkładów warunkowych $f(x \mid y)$ i $f(y \mid x)$

"Gibbs sequence": $Y'_0, X'_0, Y'_1, X'_1, Y'_2, X'_2, \ldots, Y'_k, X'_k$

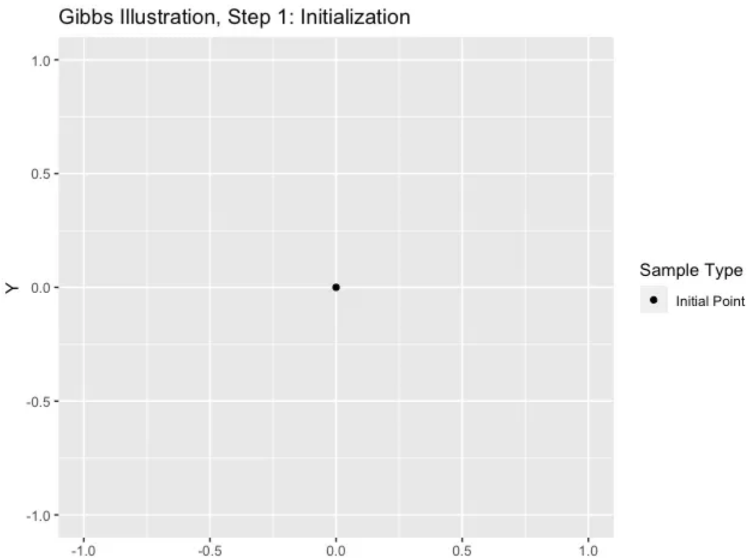
Początkowa wartość $Y'_0 = y'_0$, następne elementy generowane według schematu: [2]

$$X'_j \sim f(x|Y'_j = y'_j)$$

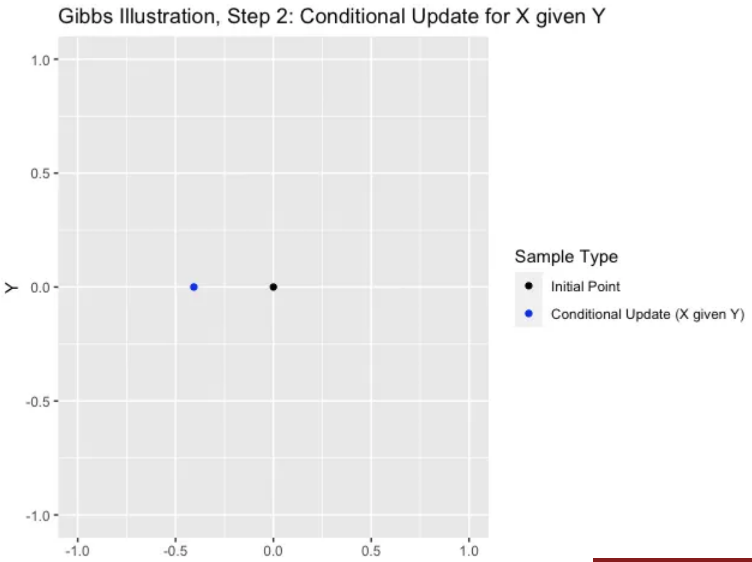
$$Y'_{j+1} \sim f(y|X'_j = x'_j)$$

Konkretniej

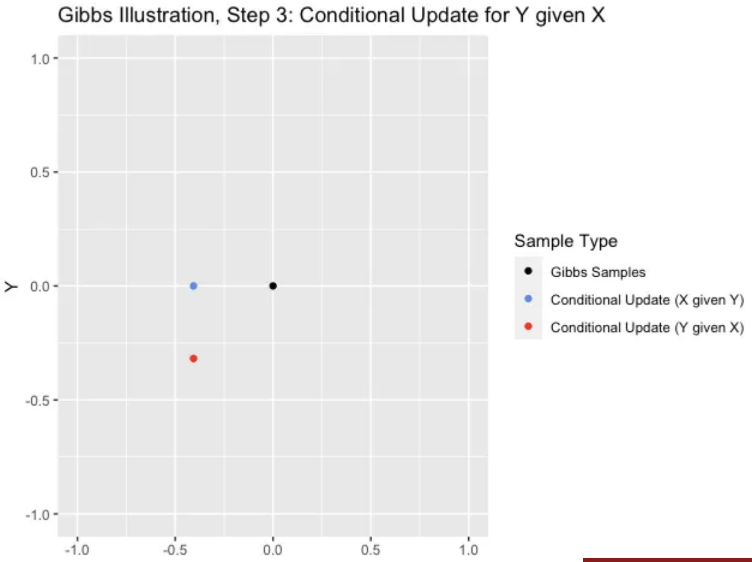
$$X \mid (Y = y) \sim N(\rho y, 1 - \rho^2)$$
$$Y \mid (X = x) \sim N(\rho x, 1 - \rho^2)$$ [1]

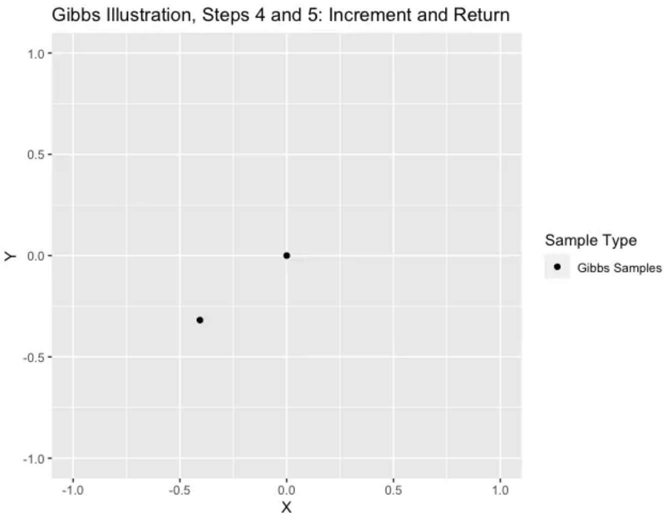

Gibbs Illustration, Step 1: Initialization
1.0
0.5
0.0
-0.5
-1.0
Y
-1.0
-0.5
0.0
0.5
1.0
Sample Type
Initial Point

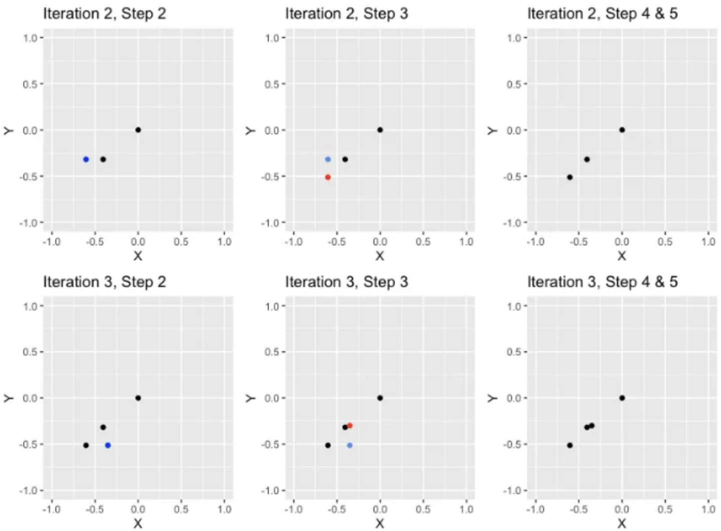
$$X_1 \mid (Y_0 = 0) \sim N(0 \cdot \rho, 1 - \rho^2)$$

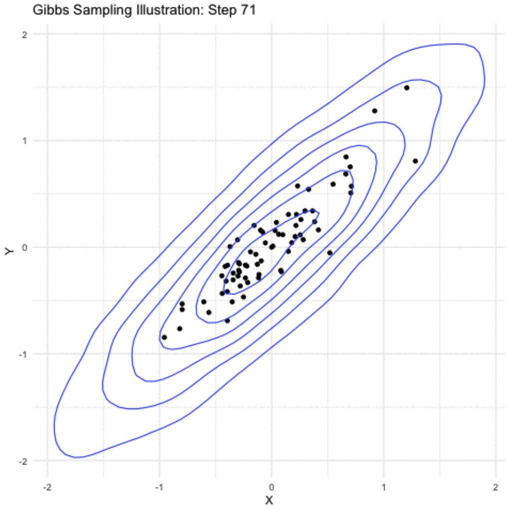

Gibbs Illustration, Step 2: Conditional Update for X given Y
1.0
0.5
0.0
-0.5
-1.0
Y
-1.0
-0.5
0.0
0.5
1.0
Sample Type
Initial Point
Conditional Update (X given Y)

$$Y_1 \mid (X_1 = -0.4) \sim N(-0.4 \cdot \rho, 1 - \rho^2)$$


Gibbs Illustration, Step 3: Conditional Update for Y given X
1.0
0.5
Y 0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0
Sample Type
Gibbs Samples
Conditional Update (X given Y)
Conditional Update (Y given X)


Gibbs Illustration, Steps 4 and 5: Increment and Return
1.0
0.5
0.0
-0.5
-1.0
Y
-1.0
-0.5
0.0
0.5
1.0
X
Sample Type
Gibbs Samples


Iteration 2, Step 2
Iteration 2, Step 3
Iteration 2, Step 4 & 5
Iteration 3, Step 2
Iteration 3, Step 3
Iteration 3, Step 4 & 5
Y
X
1.0
0.5
0.0
-0.5
-1.0


Gibbs Sampling Illustration: Step 71
Y
X
2
1
0
-1
-2
-2
-1
0
1
2

Wizualizacja algorytmu

https://chi-feng.github.io/mcmc-demo/app.html?algorithm=GibbsSampling&target=banana

Rzut monetą :)

Przeprowadzamy k eksperymentów. W każdym:

- Rzucamy monetą n razy (n jest niewiadomą)
- Prawdopodobieństwo wyrzucenia reszki oznaczmy jako θ (niewiadoma)
- Jako X_i oznaczmy liczbę wyrzuconych reszek

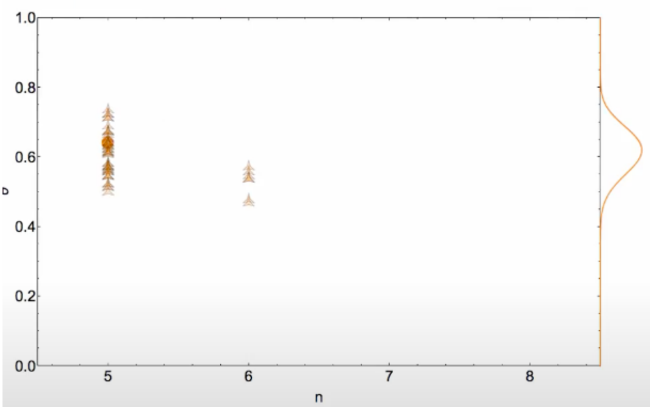
Otrzymamy dzięki temu wektor $(X_1, X_2, \ldots, X_k)$

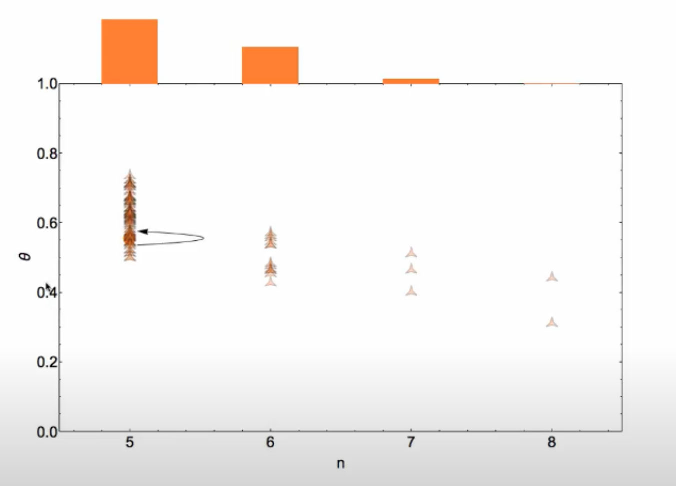
$\overline{X} = (X_1, X_2, \ldots, X_k)$
$\theta \sim U(0, 1)$ $n \sim U(5, 8)$
$P(n, \theta \mid \overline{X}) \propto P(\overline{X} \mid n, \theta) \times P(n, \theta) =$

$$= \prod_{i=1}^{k} \binom{n}{X_i} \theta^{X_i} (1-\theta)^{n-X_i} = \cdots = \theta^{kX} (1-\theta)^{nk-kX} \prod_{i=1}^{k} \binom{n}{X_i}$$

$$(\theta \mid n, \overline{X}) \sim \theta^{kX} (1-\theta)^{nk-kX} \equiv beta(kX+1, k(n-\overline{X})+1)$$

$$(n \mid \theta, \overline{X}) \sim (1-\theta)^{nk} \prod_{i=1}^{k} \binom{n}{X_i}$$


1.0
0.8
0.6
0.4
0.2
0.0
5
6
7
8
n


1.0
0.8
0.6
0.4
0.2
0.0
θ
5
6
7
8
n

Gibbs Sampling with People

- Najistotniejszym problemem w kogniwistyce i machine learningu jest zrozumienie jak ludzie uzyskują reprezentację semantyczną z obiektów percepcyjnych. [3]
- Markov Chain Monte Carlo with People (MCMCP) jest ważną metodą analizy tych zależności, asymptotycznie jest zadowalający, ale opiera się na binarnych wyborach.
- Okazuje się, że lepszy jest Gibbs Sampling with People (GSP) oparty na ciągłych suwaczkach.
- Eksperymenty sprawdzające GSP w czterech dziedzinach: kolorów, akordów muzycznych, emocji wynikających z głosu i mimiki twarzy

Algorytm

Niech $p(z_1, z_2, \ldots, z_n)$ będzie docelowym, n-wymiarowym rozkładem, z którego chcemy próbkować.
Algorytm:

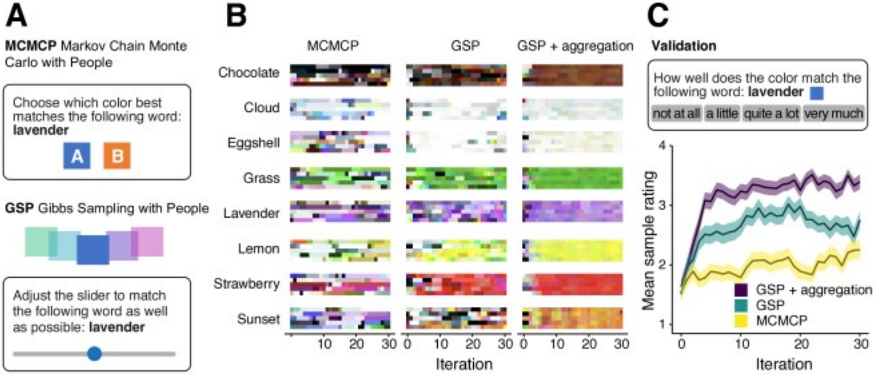
1. Wybierz początkowy wektor stanu $z^{(1)} = (z_{(1)1}, \ldots, z_{(1)n})$
2. Aktualizuj współrzędne za pomocą próbkowania z

$$p(z_{(i+1)k} \mid z_{(i+1)1}, \ldots, z_{(i+1)k-1}, z_{(i)k+1}, \ldots, z_{(i)n})$$

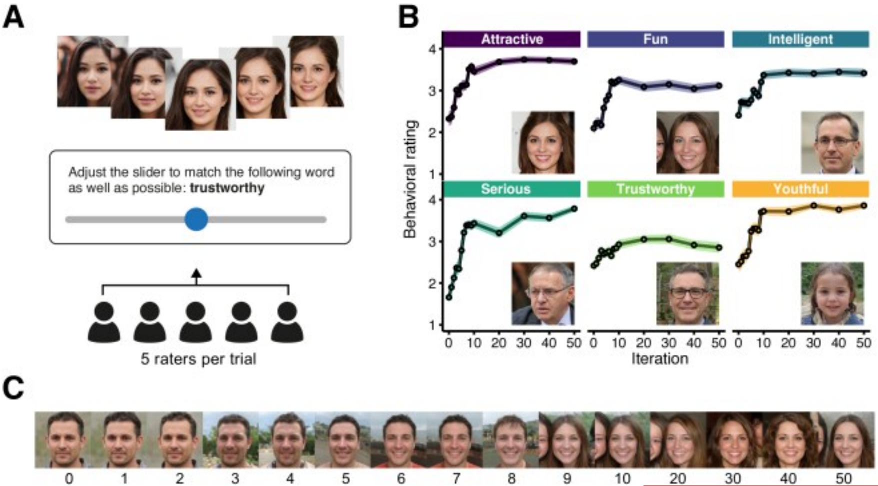
W naszym przypadku to uczestnik zapewnia próbkę z warunkowego rozkładu. Osiągane jest to za pomocą suwaka związanego z aktualnym wymiarem bodźca z_k i przesunięcie go tak, aby jak najbardziej pasował do danego zagadnienia, np. jak bardzo przyjemny jest dany dźwięk czy podobieństwo wyglądu owocu do truskawki.

Kolory

Sparametryzowana przestrzeń kolorów za pomocą schematu HSL (Hue, Saturation, Lightness) o wartościach z zakresu odpowiednio [0,360], [0,100], [0,100]



Twarze


A
Adjust the slider to match the following word as well as possible: trustworthy
5 raters per trial
B
Attractive
Fun
Intelligent
Serious
Trustworthy
Youthful
Behavioral rating
Iteration
C
0 1 2 3 4 5 6 7 8 9 10 20 30 40 50

Rozkłady warunkowe nie zawsze poprawnie wyznaczają rozkłady brzegowe

Weźmy X i Y, które mają rozkłady brzegowe:

$$f(x \mid y) \propto ye^{-yx}, \text{ gdzie } 0 < x < \infty$$
$$g(y \mid x) \propto xe^{-xy}, \text{ gdzie } 0 < y < \infty$$

Równanie całkowego punktu stałego (fixed point integral equation):

$$f_X(x) = \int \left[\int f_{X|Y}(x \mid y) f_{Y|X}(y \mid t) dy \right] f_X(t) dt$$

którego rozwiązaniem jest $f_X(x)$

$$f_X(x) = \int \left[\int ye^{-yx} te^{-ty} dy \right] f_X(t)dt = \int \left[\frac{t}{(x+t)^2} \right] f_X(t)dt$$

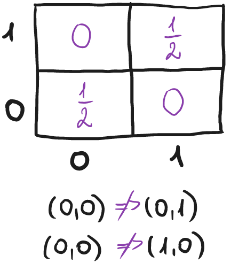
podstawiając $f_X(t) = \frac{1}{t}$:

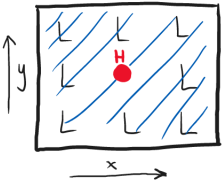
$$\frac{1}{x} = \int \left[\frac{t}{x+t} \right] \frac{1}{t} dt$$

Rozwiązaniem jest zatem $\frac{1}{x}$, jednak nie jest to funkcja gęstości.
Ważnym zatem założeniem jest, aby $\int f_X(x)dx < \infty$

Wady

- Długi okres zbieżności, w szczególności przy przestrzeniach o wyższych wymiarach. Ma na to wpływ także kształt rozkładu.
- Początkowe próbki zazwyczaj nie odwzorowują dokładnie oczekiwanego rozkładu, dlatego też często są pomijane.





Bibliografia

Seth Billiau.
Gibbs sampling explained, May 2021.

George Casella and Edward I George.
Explaining the gibbs sampler.
The American Statistician, 46(3):167–174, 1992.

Peter M. C. Harrison.
Gibbs sampling with people, Aug 2020.

DZIĘKUJĘ ZA UWAGĘ :)