２次元ガウス分布でギブスサンプリング

jojonki

May 2019

1　はじめに

この文書は@jojonki がギブスサンプリングを勉強する過程で色々調べたので，その調べた内容を実例を持って説明するものである．具体的な例としては，２次元ガウス分布に対してギブスサンプリングを適用して，２変量のガウス分布に沿うものを条件付き確率分布から得るところまで，数式展開及び Python のプログラムを持って説明する．また前もって言っておきたいが，私は初心者であるからして，あくまで参考程度に見てほしい．また文末の参考情報に載せていただいたものと基本的に同じものである（数式展開に関しては本文書ではかなり丁寧にやったつもり）．誤り箇所があれば Twitter 等で@jojonki までメンションを投げてもらえると助かる．

2　ギブスサンプリングとは何か？

まずサンプリングとは，解析的に計算できない確率分布を知るための手段である．大量にサンプリングを行い，その平均を取ることで知りたい確率分布を近似する手法である．

サンプリングを解説する文献は色々多いのが，最初にギブスサンプリングを説明している文献が多い気がするので，私もこのギブスサンプリングから勉強をしている．ギブスサンプリングを誤解を恐れて言うと，得たい確率分布を条件付き確率分布に分解し，１確率変量ごとにマルコフ過程に従ってサンプリングするというものである．一般に，複数の確率変数がある場合の同時分布からサンプルするのは難しいので，条件付き確率を変数の番号 i をランダム（または規則的）に選び，i 以外の要素を固定した条件付き確率でを利用することでサンプルを行う．ギブスサンプリング使う条件として，この条件付き確率が計算できる必要がある．

たとえば $\boldsymbol{z} = (z_1, z_2, z_3)$ のような３変量の確率変数の確率分布からのサンプリングを行うためには，$p(z_1, z_2, z_3)$ からサンプルする必要があるが，この分布から直接サンプルを行う代わりに，１つを除いて確率変量を固定して条件付け確率を解いていく．具体的に書くと下記のような条件付き確率になり，これを繰り返し更新していくことで元の同時分布に近いものを得てやろう，というのがギブスサンプリング．このときの３変量の初期値はたぶん適当で良い，なぜなら最初の方にサンプルされた値はバーンインという区間に入っているのでどちらにしろ捨てられる事が前提とされる．

$$
\begin{aligned}
z_1 &\sim p(z_1|z_2, z_3) \\
z_2 &\sim p(z_2|z_1, z_3) \\
z_3 &\sim p(z_3|z_1, z_2)
\end{aligned}
$$

3　２次元ガウス分布に対して，ギブスサンプリングを適用する

先程の章では，かなりざっくりとギブスサンプリングを説明したので，この章では具体的な事例を持ってギブスサンプリングを説明する．ここでは，２次元ガウス分布を真の分布としたときに，ギブスサンプリングを使うことで真の分布に近しい分布を確立することを目標としよう．

まず多次元ガウス分布（D 次元）の定義を思い出そう．

$$
\mathbf{N}(\boldsymbol{x}|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^D|\Sigma|}} \exp\Big\{ -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big\} \tag{1}
$$

D 次元は辛いので，2 次元ガウス分布を考える．このときのパラメタは何でも良いと思うのだが，「続・わかりやすい　パターン認識」のギブスサンプリングで使われているパラメタを持ってくる．

$$\boldsymbol{\mu} = \mathbf{0} \tag{2}$$

$$\Sigma^{-1} = \begin{pmatrix} 1 & -a \\ -a & 1 \end{pmatrix} \tag{3}$$

a は $-1 < a < 1$ を満たす．

共分散行列部分を後で使うので，逆行列及びその行列式を計算しておく．

$$\Sigma = \frac{1}{(1-a^2)} \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} \tag{4}$$

$$|\Sigma| = \frac{1}{1-a^2} \tag{5}$$

今求めたい 2 つの変量の同時分布は下記のように展開できる．

$$\begin{aligned}
p(\boldsymbol{x}) &= p(x_1, x_2) && (6)\\
&= \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\left\{ -\frac{1}{2}(x_1, x_2)\Sigma^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\} && (7)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(x_1, x_2) \begin{pmatrix} 1 & -a \\ -a & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\} && (8)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(x_1 - ax_2, -ax_1 + x2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\} && (9)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(x_1^2 - ax_1x_2 - ax_1x_2 + x_2^2) \right\} && (10)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(x_1^2 - 2ax_1x_2 + x_2^2) \right\} && (11)
\end{aligned}$$

ここで x_1 で平方完成する（x_2 でも良い）．

$$\begin{aligned}
p(x_1, x_2) &= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}\left((x_1 - ax_2)^2 - a^2x_2^2 + x_2^s\right) \right\} && (12)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}\left((1-a^2)x_2^2 + (x_1 - ax_2)^2\right) \right\} && (13)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \exp\left\{ -\frac{1}{2}(x_1 - ax_2)^2) \right\} && (14)
\end{aligned}$$

ここからのサンプリングは難しいと仮定するとき，ギブスサンプリングでは条件付き確率を求めるのであった．つまり，$p(x_1|x_2)$ と $p(x_2|x_1)$ からサンプリングしたい．まず $p(x_1|x_2)$ をベイズの定理を使って導出してみる．

$$\begin{aligned}
p(x_1|x_2) &= \frac{p(x_1, x_2)}{p(x_2)} && (15)\\
&= \frac{p(x_1, x_2)}{\int p(x_1, x_2)dx_1} && (16)
\end{aligned}$$

先程計算した同時確率が使えそうである．分子はそのまんまなのだが，分母は x_1 で周辺化されているので，ガウス積分の公式を利用するとキレイな式になる．分母に注目して計算していこう．

$$\begin{aligned}
\int p(x_1, x_2)dx_1 &= \int \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \exp\left\{ -\frac{1}{2}(x_1 - ax_2)^2) \right\} dx_1 && (17)\\
&= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \int \exp\left\{ -\frac{1}{2}(x_1 - ax_2)^2) \right\} dx_1 && (18)
\end{aligned}$$

x_1 に関する積分であるため，最後の項にまで積分エリアを絞れる．次に $t = x_1 - ax_2$ と置く．x_2 は $p(x_1|x_2)$ 環境下では何らかの値が観測されているので確率値ではない．また a も定数であるので，t が x_1 に関する関数になっていることは変わらないので問題がない．

$$\int p(x_1, x_2)dx_1 = \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \int \exp(-\frac{1}{2}t^2))dt \tag{19}$$

ここでガウス積分の公式 $\int \exp(-ax^2)dx = \sqrt{\frac{\pi}{a}}$ を利用すると，

$$\begin{aligned} \int p(x_1, x_2)dx_1 &= \frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \sqrt{2\pi} &(20) \\ &= \sqrt{\frac{1-a^2}{2\pi}} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} &(21) \\ & &(22) \end{aligned}$$

これを式 (16) に代入すると，

$$\begin{aligned} p(x_1|x_2) &= \frac{p(x_1, x_2)}{\int p(x_1, x_2)dx_1} &(23) \\ &= \frac{\frac{\sqrt{1-a^2}}{2\pi} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\} \exp\left\{ -\frac{1}{2}(x_1 - ax_2)^2) \right\}}{\sqrt{\frac{1-a^2}{2\pi}} \exp\left\{ -\frac{1}{2}(1-a^2)x_2^2 \right\}} &(24) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}(x_1 - ax_2)^2) \right\} &(25) \end{aligned}$$

と，かなりキレイになる．そして冷静に見るとこれは $\mu = ax_2$，$\sigma^2 = 1$ の 1 次元ガウス分布になっている．2 次元ガウス分布も 1 次元を観測した状態の条件付き確率は，1 次元ガウス分布と次元が下がるわけである．これでサンプリング計算できるようになった．

ところで式 (11) から分かるように，x_1 と x_2 は入れ替えても同じの対称性があることが分かるので，$p(x_2|x_1)$ は先程の結果かから，x_1 と x_2 を入れ替えるだけで出来上がる．

$$p(x_2|x_1) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}(x_2 - ax_1)^2) \right\} \tag{26}$$

4 Python で書いてみる

TeX 上でのコードは見にくいと思うので，`https://github.com/jojonki/gibbs-sampling` に Jupyter Notebook も用意したのでそちらを見ても良い．

まず求めたい 2 次元ガウス分布（真の分布）をプロットする．

```python
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import animation

a = -0.5

# print a true gauss distribution
def gauss(x, y, a):
    def _gauss(x, y):
        """Multivariate normal distribution (two variables)"""
        mu = np.array([0, 0])
        sigma = (1 - a*a) * np.array([[1 , a],
                                      [a,  1]])
        det = np.linalg.det(sigma)
        inv_sigma = np.linalg.inv(sigma)

        x_c = np.array([x, y]) - mu
        return np.exp(- x_c.dot(inv_sigma).dot(x_c[np.newaxis, :].T)/2.0) / (2*np.pi*np.sqrt
            (det))
```

```
    z = np.vectorize(_gauss)(x, y)
    return z

def plot_gauss():
    t = np.arange(-5, 5, 0.1)
    X, Y =np.meshgrid(t, t)
    Z = gauss(X, Y, a)

    plt.title('True Distribution(a={})'.format(a))
    c = plt.contour(X, Y, Z)
    c.clabel(fmt='%1.2f', fontsize=10)

plot_gauss()
plt.show()
```

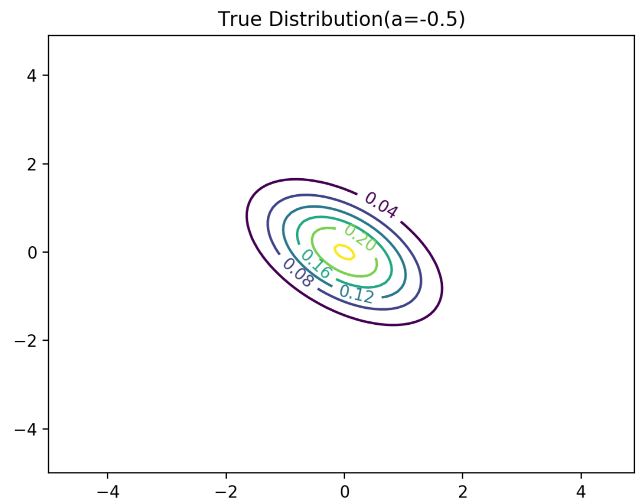

Figure 1: 求めたい 2 次元ガウス分布

それではギブスサンプリングを行っていく．適当な初期値に対して x_1 及び x_2 を交互に固定して，1 次元ガウス分布から値をサンプルすれば良い．

```
def gibbs_sampling(a, step, burn_in):
    x = np.zeros(2) # initial x
    samples = np.array(x)
    for i in range(step):
        denomi = 1 / np.sqrt(2*np.pi)
        x[0] = denomi * np.random.normal(a*x[1], 1) # mu=ax[1], sigma=1
        samples = np.append(samples, (x))
        x[1] = denomi * np.random.normal(a*x[0], 1)
        samples = np.append(samples, (x))
    samples = samples.reshape((2*step+1, x.shape[0])) # +1 means initial x

    return samples


def onetime_gibbs_sampling(_):
    denomi = 1 / np.sqrt(2*np.pi)
    x[0] = denomi * np.random.normal(a*x[1], 1) # mu=ax[1], sigma=1
    x[1] = denomi * np.random.normal(a*x[0], 1)

    plt.scatter(x[0], x[1], s=10, c='pink', edgecolor='red')
    plt.ylim(-5,5)
    plt.xlim(-5,5)
```

```
anim_plot = False
if anim_plot:
    fig = plt.figure()
    x = np.zeros(2)
    ani = animation.FuncAnimation(fig, onetime_gibbs_sampling, interval=10, blit=True)
else:
    step = 3000
    burn_in = 1000
    samples = gibbs_sampling(a, step, burn_in)
    plt.scatter(samples[burn_in:, 0], samples[burn_in:, 1], s=10, c='pink', alpha=0.2,
        edgecolor='red')

# ani.save('gibbs_sampling.gif', writer='imagemagick')
plot_gauss()
plt.show()
```

ギブスサンプリングした結果を赤い点で下記に示す．なんとなくうまくいっている気がする．プログラム内の anim_plot を True にすれば，アニメーションでサンプリングの様子が確認できる．

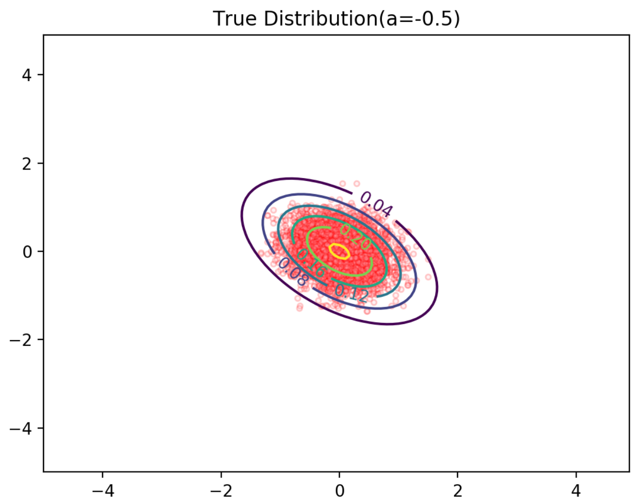

Figure 2: ギブスサンプリングした結果

5 まとめ

ギブスサンプリングを 2 次元ガウス分布を土台に説明してみた．とりあえずこの手法がどれぐらい強いのか，もっと例を知らないとよくわからないが，まず第一弾としては上々であろう．

6 参考情報

ギブスサンプリングに関して参考にさせていただいたところ

- 2 次元正規分布でギブスサンプリングする【Python】
  https://qiita.com/ManInML/items/2a43028ce9cca2c55905
- 第 11 章 サンプリング法. 博士課程 1 年. 原 祐輔 http://bin.t.u-tokyo.ac.jp/prml2009/ppt/PRML11.ppt

ガウス分布やガウス積分公式についてはこちら．

- 一次元の正規分布から多次元正規分布へ，HELLO CYBERNETICS
  https://www.hellocybernetics.tech/entry/2016/10/06/111153
- ガウス積分公式集（証明付き）
  https://risalc.info/src/gaussian-integral.html#gen