指数型分布族と双対座標系 Exponential Family and dual coordinate systems

ryamada

2017年3月6日

1 はじめに Introduction

- 正規分布を例に都合のよい２つの座標系があることを示し、それらは平坦で、双直交であると説明した。

$$\theta_1 = \frac{m}{s^2}$$
$$\theta_2 = -\frac{1}{2s^2}$$
$$m = -\frac{\theta_1}{2\theta_2}$$
$$s^2 = -\frac{1}{2\theta_2}$$
$$\eta_1 = m$$
$$\eta_2 = m^2 + s^2$$
$$m = \eta_1$$
$$s^2 = \eta_2 - \eta_1^2$$

- この座標系のとり方を説明するために、分布の指数型表現というものを使う。

2 正規分布の指数型表現

$$P(x|m,s) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(x-m)^2}{2s^2}}$$

- 対数を取り、xとパラメタが関係する部分と、パラメタのみにかかわる部分(とxのみにかかわる部分(以下の例ではない))とに分ける。

$$\log P(x|m,s) = -\frac{(x-m)^2}{2s^2} - \frac{1}{2}\log 2\pi s^2$$
$$= \frac{m}{s^2}x - \frac{1}{2s^2}x^2 - (\frac{m^2}{2s^2} + \frac{1}{2}\log 2\pi s^2)$$
$$= \left(\frac{m}{s^2}, -\frac{1}{2s^2} \right) \cdot \begin{pmatrix} x \\ x^2 \end{pmatrix} - (\frac{m^2}{2s^2} + \frac{1}{2}\log 2\pi s^2)$$

- 次のように書き表すことにする。

$$\log P(x|\theta_1,\theta_2) = \left(\theta_1, \theta_2 \right) \cdot \begin{pmatrix} x \\ x^2 \end{pmatrix} - A(\theta_1,\theta_2)$$
$$= \left(\theta_1, \theta_2 \right) \cdot \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} - A(\theta_1,\theta_2)$$
$$\theta_1 = \frac{m}{s^2}$$
$$\theta_2 = -\frac{1}{2s^2}$$
$$A(\theta_1,\theta_2) = -\frac{\theta_1^2}{4\theta_2} + \frac{1}{2}\log\frac{\pi}{\theta_2}$$
$$f_1(x) = x$$
$$f_2(x) = x^2$$

- 以下のように、指数関数の形で表されるので、「指数型」と呼ばれる。

$$P(x|\theta_1,\theta_2) = e^{\left(\theta_1, \theta_2 \right) \cdot \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} - A(\theta_1,\theta_2)}$$

- 平坦で双対座標系の片方となるθ座標系は、分布関数を指数型で表したときに、xの関数($f_i(x)$)の係数であることが知られている
- 他方、η座標系は、θ座標系と双対な関係にある座標系であるが、どのようなものがそれに相当するかというと、xの関数として現れた$f_i(x)$の期待値であることが知られている。
- 正規分布の場合には$f_1(x) = x$と$f_2(x) = x^2$のそれぞれの期待値である。平均m、標準偏差sの正規分布のxの期待値はmそのものであるし、$(x-m)^2$の期待値がs^2であるから、x^2の期待値は、$m^2 + s^2$である。

3 η座標系

- 指数型表現をしたときのxの関数の係数がθ座標であり、xの関数の期待値がη座標である。
- η座標は次のようにも表せる。分布関数のxによらない成分、θのみによる成分に「双対対応」するのがηであり、その「双対対応」というのは偏微分をとることである、というように読める。

$$\eta_i = \frac{\partial A(\theta)}{\partial \theta_i}$$

- これは、対数尤度関数をθ_iで偏微分したものの期待値が0であることを利用することで、以下のように示せる。

$$\frac{\partial \log p}{\partial \theta_i} = f_i(x) - \frac{\partial A(\theta)}{\partial \theta_i}$$
$$\int \frac{\partial \log p}{\partial \theta_i} \times pdx = \int (f_i(x) - \frac{\partial A(\theta)}{\partial \theta_i}) \times pdx$$
$$\int \frac{\partial \log p}{\partial \theta_i} \times pdx = \int f_i(x) \times pdx - \frac{\partial A(\theta)}{\partial \theta_i} \int pdx$$
$$\int \frac{1}{p}\frac{\partial p}{\partial \theta_i} \times pdx = E[f_i(x)] - \frac{\partial A(\theta)}{\partial \theta_i} \times 1$$
$$\int \frac{\partial p}{\partial \theta_i} dx = E[f_i(x)] - \frac{\partial A(\theta)}{\partial \theta_i}$$
$$0 = E[f_i(x)] - \frac{\partial A(\theta)}{\partial \theta_i}$$

3.1 練習問題

3.1.1 指数型分布族の例 Examples of exponential family

非常に多くの理論確率分布が含まれる。どのような分布が含まれるか確認せよ。https://en.wikipedia.org/wiki/Exponential_family

There are very many theoretical probability functions in exponential family. See the URL above.

3.1.2 Canonical formの導出 Derive canonical form

2つの分布をhttps://en.wikipedia.org/wiki/Exponential_family　より選び、その指数型表現を通常のパラメタ表現から導け。

Select two distributions from the site and derive their exponential form.

3.1.3 ヤコビ行列 Jacobian matrix

正規分布のパラメタ変換におけるヤコビ行列を求めよ。 We changed the parameters of normal distribution. Show its Jacobian matrix.

$$\theta_1 = \frac{\mu}{\sigma^2}$$
$$\theta_2 = -\frac{1}{2\sigma^2}$$