

7.2.3 に関して、確認する。

(7.36) の確率分布を仮定する。累積密度関数は一般には確率分布になっていないが、 y_n が $\{-1, 1\}$ なので、確率分布になる。

$p(Y|F, X) = p(Y|F) = \prod_{n=1}^N p(y_n|f_n)$ を仮定すると、 $p(Y|X)$ を定数として、

$$p(F|Y, X)p(Y|X) = p(Y|F, X)p(F|X) = p(F|X) \prod_{n=1}^N p(y_n|f_n) \quad (1)$$

となり、(7.37) が成り立つ。

パラメータはそれぞれの $\tilde{\mu}_n, \tilde{\sigma}_n^2$ とすると、一部定数項があるとして、(4.84) を参考に、(7.38) のように仮定する。このとき、(4.82) と (4.84) 式との対比を考えると、 $f_n(\theta) = p(y_n|f_n)$ 、 $\tilde{f}_n(\theta) = t(f_n|\tilde{\mu}_n, \tilde{\sigma}_n^2)$ となる。一部、f の記号が混在しているので注意のこと。t は (7.39) のようにガウス分布として、p.199 の下から 2 行を考慮すると $p(F|X) = \mathcal{N}(F|0, K_\beta)$ になり、(4.85) と同様に平方展開を行うと、(7.40)-(7.42) が求まる。なお、F と f_n の関係を考えて、 $p(Y|F)$ は $\tilde{\mu}, \tilde{\Sigma}$ を持つ多次元ガウス分布となる。

この分布でうまく置換し、 f_i が左上になるようにする。すると、(3.31) で $x_i = \{f_i\}$ 、残りの要素が x_2 とみなせ、(3.31) より、(7.43) の μ_i, σ_i^2 が計算でき、(7.43) のパラメータの初期値がもとまる。(4.86) のように \tilde{f}_i に相当する、 t で割る。(4.85) のように平方完成すると、(7.44)-(7.46) が求まる。

次に、(4.87) のように、f に相当するものをかけて、r を求める。

その後、(4.65),(4.66) から、(7.43) のパラメータを更新する。(4.65),(4.66) の計算には Z の微分が必要になる。(7.47) は

$$r(f_i) = \frac{1}{Z_i} \Phi(y_i|f_i) \mathcal{N}(f_i|\mu_{\setminus i}, \sigma_{\setminus i}^2) \quad (2)$$

なので、 Z_i は (4.78)(4.79) や A.3 を参考にすると、(4.79)(4.80),(A.44) のように、

$$Z_i = \Phi(a_i)(a_i = \frac{y_n \mu_{\setminus i}}{\sqrt{1 + y_i^2 \sigma_{\setminus i}^2}} = \frac{y_n \mu_{\setminus i}}{\sqrt{1 + \sigma_{\setminus i}^2}}) \quad (3)$$

(4.65) を考慮すると、平均の更新は μ_i が $\mu_{\setminus i}$ に、 μ_{i+1} が $\hat{\mu}_i$ に対応するので、

$$\begin{aligned} \hat{\mu}_i &= \mu_{\setminus i} + \sigma_{\setminus i}^2 \frac{\partial}{\partial \mu_{\setminus i}} \ln Z_i = \mu_{\setminus i} + \sigma_{\setminus i}^2 \frac{\partial}{\partial Z_i} \ln Z_i \frac{\partial Z_i}{\partial a_i} \frac{\partial a_i}{\partial \mu_{\setminus i}} = \mu_{\setminus i} + \sigma_{\setminus i}^2 \frac{1}{Z_i} \frac{\partial}{\partial a_i} \int_{-\infty}^{a_i} \mathcal{N}(x|0, 1) dx \frac{y_n}{\sqrt{1 + \sigma_{\setminus i}^2}} \\ &= \mu_{\setminus i} + \sigma_{\setminus i}^2 \frac{1}{\Phi(a_i)} \mathcal{N}(a_i|0, 1) \frac{y_n}{\sqrt{1 + \sigma_{\setminus i}^2}} = \mu_{\setminus i} + \frac{y_n \sigma_{\setminus i}^2 \mathcal{N}(a_i|0, 1)}{\Phi(a_i) \sqrt{1 + \sigma_{\setminus i}^2}} \end{aligned} \quad (4)$$

となり、(7.48) が求まる。(4.66) を考慮すると、平均と同様に

$$\begin{aligned} \hat{\sigma}_i^2 &= \sigma_{\setminus i}^2 - (\sigma_{\setminus i}^2)^2 \left\{ \left(\frac{1}{\Phi(a_i)} \mathcal{N}(a_i|0, 1) \frac{y_n}{\sqrt{1 + \sigma_{\setminus i}^2}} \right)^2 - 2 \frac{\partial}{\partial Z_i} \ln Z_i \frac{\partial Z_i}{\partial a_i} \frac{\partial a_i}{\partial \sigma_{\setminus i}^2} \right\} \\ &= \sigma_{\setminus i}^2 - \sigma_{\setminus i}^4 \left\{ \frac{y_i^2 \mathcal{N}(a_i|0, 1)^2}{\Phi(a_i)^2 (1 + \sigma_{\setminus i}^2)} - 2 \frac{\partial}{\partial Z_i} \ln Z_i \frac{\partial Z_i}{\partial a_i} \frac{\partial a_i}{\partial \sigma_{\setminus i}^2} \right\} \\ &= \sigma_{\setminus i}^2 - \sigma_{\setminus i}^4 \left\{ \frac{y_i^2 \mathcal{N}(a_i|0, 1)^2}{\Phi(a_i)^2 (1 + \sigma_{\setminus i}^2)} - 2 \frac{1}{\Phi(a_i)} \mathcal{N}(a_i|0, 1) \left(-\frac{y_i \mu_{\setminus i}}{2(\sqrt{1 + \sigma_{\setminus i}^2})^3} \right) \right\} \\ &= \sigma_{\setminus i}^2 - \frac{\sigma_{\setminus i}^4 \mathcal{N}(a_i|0, 1)}{\Phi(a_i)(1 + \sigma_{\setminus i}^2)} \left\{ \frac{y_i^2 \mathcal{N}(a_i|0, 1)}{\Phi(a_i)} + a_i \right\} = \sigma_{\setminus i}^2 - \frac{\sigma_{\setminus i}^4 \mathcal{N}(a_i|0, 1)}{\Phi(a_i)(1 + \sigma_{\setminus i}^2)} \left\{ a_i + \frac{\mathcal{N}(a_i|0, 1)}{\Phi(a_i)} \right\} \end{aligned} \quad (5)$$

最後に $y_i = \{-1, 1\}$ に注意する。