

7.3 に関して、式を見直してみる。

(7.54),(7.55) にあるように

$$p(F_*, F_X, F_Z|Y) = p(F_{all}|Y) \approx q(F_{all}) \equiv p(F_*, F_X|F_Z)q(F_Z) \quad (1)$$

右辺に関して、Y の条件付がなくなっているような形で近似する。これは一般的な変分推論のように (7.56) の KL ダイバージェンスを最小化する。

KL ダイバージェンスに関して、それに含まれる $p(F_{all}|Y)$ を (7.57) のように考える。ガウス過程では Y は F_X のみに依存しているので、(あえて書くと Y_X のようにかけて、 Y_Z, Y_* に相当する部分は未観測。)

$$\begin{aligned} p(F_{all}|Y)p(Y) &= p(F_{all}, Y) = p(F_*, F_X, F_Z, Y) = p(Y|F_*, F_X, F_Z)p(F_*|F_X, F_Z)p(F_X, F_Z) \\ &= p(Y|F_X)p(F_*|F_X, F_Z)p(F_X, F_Z) \end{aligned} \quad (2)$$

となり、(7.57) のように

$$p(F_{all}|Y) = \frac{p(Y|F_X)p(F_*|F_X, F_Z)p(F_X, F_Z)}{p(Y)} \quad (3)$$

が求まる。

これと (7.55) の近似を (7.56) に代入する。

$$\begin{aligned} D_{KL}[q(F_{all})||p(F_{all}|Y)] &= - \int q(F_{all}) \ln \frac{p(F_{all}|Y)}{q(F_{all})} dF_{all} \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(Y|F_X)p(F_*|F_X, F_Z)p(F_X, F_Z)}{p(F_*, F_X|F_Z)q(F_Z)p(Y)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(Y|F_X, F_Z)p(F_*|F_X, F_Z)p(F_X, F_Z)}{p(F_*, F_X|F_Z)q(F_Z)p(Y)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(Y, F_X, F_Z)p(F_*|F_X, F_Z)}{p(F_*, F_X|F_Z)q(F_Z)p(Y)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(F_X, F_Z|Y)p(F_*|F_X, F_Z)}{p(F_*, F_X|F_Z)q(F_Z)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(F_X, F_Z|Y)p(F_*|F_X, F_Z)}{p(F_*|F_X, F_Z)p(F_X|F_Z)q(F_Z)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)q(F_Z) \ln \frac{p(F_X, F_Z|Y)}{p(F_X|F_Z)q(F_Z)} dF_* dF_X dF_Z \\ &= - \int \int \int p(F_*, F_X|F_Z)dF_* q(F_Z) \ln \frac{p(F_X, F_Z|Y)}{p(F_X|F_Z)q(F_Z)} dF_X dF_Z \\ &= - \int \int p(F_X|F_Z)q(F_Z) \ln \frac{p(F_X, F_Z|Y)}{p(F_X|F_Z)q(F_Z)} dF_X dF_Z \end{aligned} \quad (4)$$

最後の部分は KL ダイバージェンスとみなせ、 $q(F_X, F_Z) \equiv p(F_X|F_Z, Y)q(F_Z)$ として、(7.59) が成り立つ。

$$D_{KL}[q(F_{all})||p(F_{all}|Y)] = D_{KL}[q(F_X, F_Z)||p(F_X, F_Z|Y)] \quad (5)$$

(4.29),(4.30) を踏まえると、一般に (F_{all}) でも $\{F_X, F_Z\}$ でも、(7.60),(7.61) が求まる。

$$\ln p(Y) = \mathcal{L}[q] + D_{KL}[q(F_X, F_Z)||p(F_X, F_Z|Y)] \quad (6)$$

$$\mathcal{L}[q] = \int q(\{F_X, F_Z\}) \ln \frac{p(Y, \{F_X, F_Z\})}{q(\{F_X, F_Z\})} d\{F_X, F_Z\} = \int \int q(F_X, F_Z) \ln \frac{p(Y, F_X, F_Z)}{q(F_X, F_Z)} dF_X dF_Z \quad (7)$$

$p(Y)$ が何らかの定数であることを考慮すると、(7.59) の最小化は (7.61) の最大化と等価になる。

(7.61) の最大化は付録 A.2 にあるように厳密に行える。(7.61) は $\ln G(F_Z, Y) \equiv \int p(F_X|F_Z) \ln p(Y|F_X) dF_X$ と置くと (A.25) のように変形できる。

$$\begin{aligned}
\mathcal{L}[q] &= \int \int q(F_X, F_Z) \ln \frac{p(Y, F_X, F_Z)}{q(F_X, F_Z)} dF_X dF_Z \\
&= \int \int p(F_X|F_Z) q(F_Z) \ln \frac{p(Y|F_X, F_Z) p(F_X|F_Z) p(F_Z)}{p(F_X|F_Z) q(F_Z)} dF_X dF_Z \\
&= \int \int p(F_X|F_Z) q(F_Z) \ln \frac{p(Y|F_X) p(F_X|F_Z) p(F_Z)}{p(F_X|F_Z) q(F_Z)} dF_X dF_Z \\
&= \int q(F_Z) \left\{ \int p(F_X|F_Z) \ln \frac{p(Y|F_X) p(F_Z)}{q(F_Z)} dF_X \right\} dF_Z \\
&= \int q(F_Z) \left\{ \int \left(p(F_X|F_Z) \ln p(Y|F_X) + p(F_X|F_Z) \ln \frac{p(F_Z)}{q(F_Z)} \right) dF_X \right\} dF_Z \\
&= \int q(F_Z) \left\{ \left(\int p(F_X|F_Z) \ln p(Y|F_X) dF_X + \int p(F_X|F_Z) \ln \frac{p(F_Z)}{q(F_Z)} dF_X \right) \right\} dF_Z \\
&= \int q(F_Z) \left\{ \left(\int p(F_X|F_Z) \ln p(Y|F_X) dF_X + \ln \frac{p(F_Z)}{q(F_Z)} \right) \right\} dF_Z \\
&= \int q(F_Z) \left\{ \left(\ln G(F_Z, Y) + \ln \frac{p(F_Z)}{q(F_Z)} \right) \right\} dF_Z
\end{aligned} \tag{8}$$

さて、F はガウス過程なので、X,Z,*を入力とする（ここでは平均 0 も仮定されている）ガウス分布となる。つまり、例えば、以下ようになる。

$$p(F_X, F_Z) = p(F_X, F_Z|X, Z) = \mathcal{N}([F_X, F_Z]^T | [0, 0]^T, \begin{pmatrix} K_{XX} & K_{XZ} \\ K_{ZX} & K_{ZZ} \end{pmatrix}) \tag{9}$$

$$\Lambda = \begin{pmatrix} \Lambda_{XX} & \Lambda_{XZ} \\ \Lambda_{ZX} & \Lambda_{ZZ} \end{pmatrix} \equiv \begin{pmatrix} K_{XX} & K_{XZ} \\ K_{ZX} & K_{ZZ} \end{pmatrix}^{-1} \quad \text{とすると、(A.3) より、}$$

$$\begin{aligned}
&\begin{pmatrix} \Lambda_{XX} & \Lambda_{XZ} \\ \Lambda_{ZX} & \Lambda_{ZZ} \end{pmatrix} \equiv \begin{pmatrix} K_{XX} & K_{XZ} \\ K_{ZX} & K_{ZZ} \end{pmatrix}^{-1} \\
&= \begin{pmatrix} (K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1} & -(K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1}K_{XZ}K_{ZZ}^{-1} \\ -K_{ZZ}^{-1}K_{ZX}(K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1} & K_{ZZ}^{-1} + K_{ZZ}^{-1}K_{ZX}(K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1}K_{XZ}K_{ZZ}^{-1} \end{pmatrix}
\end{aligned} \tag{10}$$

となるから、(A.7)-(A.9) より、

$$p(F_X|F_Z) = \mathcal{N}(F_X | \mu_{X|Z}, \Sigma_{X|Z}) \tag{11}$$

$$\Sigma_{X|Z} = \Lambda_{XX}^{-1} = (K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1-1} = K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX} \tag{12}$$

$$\begin{aligned}
\mu_{X|Z} &= 0 - \Lambda_{XX}^{-1} \Lambda_{XZ}(F_Z - 0) = -\Lambda_{XX}^{-1} \Lambda_{XZ} F_Z \\
&= (K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})(K_{XX} - K_{XZ}K_{ZZ}^{-1}K_{ZX})^{-1} K_{XZ}K_{ZZ}^{-1} F_Z = K_{XZ}K_{ZZ}^{-1} F_Z
\end{aligned} \tag{13}$$

(7.5) のように考え、

$$p(Y|F_X) = \mathcal{N}(Y|F_X, \sigma^2 I) \tag{14}$$

となるので、

$$\begin{aligned}
\ln G(F_Z, Y) &= \int p(F_X|F_Z) \ln p(Y|F_X) dF_X \\
&= \int \mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z}) \left(-\frac{1}{2}(Y - F_X)^T \sigma^{-2} I (Y - F_X) - \frac{1}{2}(N \ln 2\pi + \ln |\sigma^2 I|) \right) dF_X \\
&= -\frac{1}{2} \int \mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z}) (\sigma^{-2} F_X^T F_X - 2\sigma^{-2} Y^T F_X + \sigma^{-2} Y^T Y + (N \ln 2\pi + \ln |\sigma^2 I|)) dF_X \\
&= -\frac{1}{2} \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [\sigma^{-2} F_X^T F_X - 2\sigma^{-2} Y^T F_X + \sigma^{-2} Y^T Y + (N \ln 2\pi + \ln |\sigma^2 I|)] \\
&= -\frac{1}{2} (\sigma^{-2} \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [F_X^T F_X] - 2\sigma^{-2} Y^T \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [F_X] \\
&\quad + \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [\sigma^{-2} Y^T Y + N \ln 2\pi + \ln |\sigma^2 I|]) \\
&= -\frac{1}{2} (\sigma^{-2} (\mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [(F_X - \mu_{X|Z})^T (F_X - \mu_{X|Z})] + \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [F_X]^T \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [F_X]) \\
&\quad - 2\sigma^{-2} Y^T \mu_{X|Z} + \sigma^{-2} Y^T Y + N \ln 2\pi + \ln |\sigma^2 I|) \\
&= -\frac{1}{2} (\sigma^{-2} (Tr[\Sigma_{X|Z}] + \mu_{X|Z}^T \mu_{X|Z}) - 2\sigma^{-2} Y^T \mu_{X|Z} + \sigma^{-2} Y^T Y + N \ln 2\pi + \ln |\sigma^2 I|) \\
&= -\frac{1}{2} ((Y - \mu_{X|Z})^T \sigma^{-2} I (Y - \mu_{X|Z}) + N \ln 2\pi + \ln |\sigma^2 I| + \sigma^{-2} Tr[\Sigma_{X|Z}]) = \ln \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) - \frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}]
\end{aligned} \tag{15}$$

なお、 $\Sigma_{X|Z} = \mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [(F_X - \mu_{X|Z})(F_X - \mu_{X|Z})^T]$ であることを考慮すると、要素を比較して、 $\mathbb{E}_{\mathcal{N}(F_X|\mu_{X|Z}, \Sigma_{X|Z})} [(F_X - \mu_{X|Z})^T (F_X - \mu_{X|Z})] = Tr[\Sigma_{X|Z}]$ だとわかる。また、 $N = \dim(Y) = \dim(F_X) = \dim(X)$ 。

これを (A.25) に代入すると、

$$\begin{aligned}
\mathcal{L}[q] &= \int q(F_Z) \left\{ \ln \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) - \frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}] + \ln \frac{p(F_Z)}{q(F_Z)} \right\} dF_Z \\
&= \int q(F_Z) \ln \frac{\mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z)}{q(F_Z)} dF_Z - \int q(F_Z) \frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}] dF_Z - \ln Z + \ln Z \\
&= \int q(F_Z) \ln \frac{\frac{1}{Z} \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z)}{q(F_Z)} dF_Z - \frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}] + \ln Z \\
&= -D_{KL}[q(F_Z) || \frac{1}{Z} \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z)] - \frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}] + \ln Z
\end{aligned} \tag{16}$$

Z は定数で、 $\frac{1}{2} \sigma^{-2} Tr[\Sigma_{X|Z}]$ も $\Sigma_{X|Z}$ が (A.29) のように求まっているので、定数と考えられ、 $\mathcal{L}[q]$ を最大にするには、 $D_{KL}[q(F_Z) || \frac{1}{Z} \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z)]$ を最小にすれば良い。なお、 Z は、 $\mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z)$ が正規化されているとは限らないため、 Z をかけている。

KL ダイバージェンスを最小にするのは、(A.32) にあるように、以下ようになる。

$$q_{opt.}(F_Z) = \frac{1}{Z} \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z) \tag{17}$$

F_Z はガウス過程 (平均 0) であるため、 $p(F_Z) = p(F_Z|Z)\mathcal{N}(F_Z|0, K_{ZZ})$ となる。よって、

$$\begin{aligned}
q_{opt.}(F_Z) &= \frac{1}{Z} \mathcal{N}(Y|\mu_{X|Z}, \sigma^2 I) p(F_Z) = \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(\frac{1}{2}(Y - \mu_{X|Z})^T \sigma^{-2} I (Y - \mu_{X|Z})\right) \exp\left(\frac{1}{2} F_Z^T K_{ZZ}^{-1} F_Z\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} \sigma^{-2} (Y^T Y - 2Y^T \mu_{X|Z} + \mu_{X|Z}^T \mu_{X|Z}) - \frac{1}{2} F_Z^T K_{ZZ}^{-1} F_Z\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} \sigma^{-2} (Y^T Y - 2Y^T K_{XZ} K_{ZZ}^{-1} F_Z + K_{XZ} K_{ZZ}^{-1} F_Z^T K_{XZ} K_{ZZ}^{-1} F_Z) - \frac{1}{2} F_Z^T K_{ZZ}^{-1} F_Z\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} \sigma^{-2} (Y^T Y - 2Y^T K_{XZ} K_{ZZ}^{-1} F_Z + F_Z^T K_{ZZ}^{-1} K_{ZX} K_{XZ} K_{ZZ}^{-1} F_Z) - \frac{1}{2} F_Z^T K_{ZZ}^{-1} F_Z\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} (\sigma^{-2} Y^T Y - 2\sigma^{-2} Y^T K_{XZ} K_{ZZ}^{-1} F_Z + F_Z^T K_{ZZ}^{-1} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ}) K_{ZZ}^{-1} F_Z)\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} (\sigma^{-2} Y^T Y + (F_Z - \sigma^{-2} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y)^T K_{ZZ}^{-1} \right. \\
&\quad \left. (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ}) K_{ZZ}^{-1} (F_Z - \sigma^{-2} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y) \right. \\
&\quad \left. - \sigma^{-4} Y^T K_{XZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZZ} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y)\right) \\
&= \frac{1}{Z} \frac{1}{Z_{tmp}} \exp\left(-\frac{1}{2} (F_Z - \sigma^{-2} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y)^T K_{ZZ}^{-1} \right. \\
&\quad \left. (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ}) K_{ZZ}^{-1} (F_Z - \sigma^{-2} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y)\right) \\
&\quad \exp\left(-\frac{1}{2} (Y^T (\sigma^{-2} I - \sigma^{-4} K_{XZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZZ} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX}) Y)\right) \\
&= \mathcal{N}(F_Z | \sigma^{-2} K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZX} Y, K_{ZZ} (\sigma^{-2} K_{ZX} K_{XZ} + K_{ZZ})^{-1} K_{ZZ}) \\
&\quad (18)
\end{aligned}$$

K_{ZZ} が対称行列のため、 $K_{ZZ}^{-1T} = K_{ZZ}^{-1}$, $K_{XZ}^T = K_{ZX}$ に注意する。また、 $Z_{tmp} = \sqrt{(2\pi)^{N_X + N_Z} |\sigma^2 I| |K_{ZZ}|}$ 。最後は正規化するので、単なるガウス分布になる。よって、(7.64)((A.34)), (7.65)((A.35)) が求まった。