7.3.1.2 を確認する。

論文 [115] の 2 章を見てみる。

(4) 式にあるように一般的な点を $\{x, y, (f)\}$ を考えたときにガウス過程では任意の f たちの組み合わせが平均 0, 分散 K のガウス分布になるので $\{F_Z, f, Z, x\}$ を考えると、(A.7),(A.3) を考慮して、(7.9) と同様に

$$p(f|x, Z, F_Z) = \mathcal{N}(f|k_{xZ}{K_{ZZ}}^{-1}F_Z, k_{xx} - k_{xZ}{K_{ZZ}}^{-1}k_{Zx}) \tag{1}$$

となる。

一般に

$$p(y|f) = \mathcal{N}(y|f, \sigma^2) \tag{2}$$

なので、本の (A.24) を踏まえると、

$$p(y|x, Z, F_Z) = \mathcal{N}(y|k_{xZ}{K_{ZZ}}^{-1}F_Z, k_{xx} - k_{xZ}{K_{ZZ}}^{-1}k_{Zx} + \sigma^2) \tag{3}$$

となり、[115] の (4) 式になる。

さて、(7.63) も考慮すると、

$$\begin{aligned}
p(y_*|Y) = p(y_*|x_*, X, Y) = p(y_*|x*, X, Y, Z) = \int p(y_*, F_Z|x*, X, Y, Z)dF_Z \\
= \int p(y_*|F_Z, x*, X, Y, Z)p(F_Z|x*, X, Y, Z)dF_Z \approx \int p(y_*|F_Z, x*, Z)p(F_Z|x*, X, Y, Z)dF_Z \\
= \int p(y_*|F_Z, x*, Z)p(F_Z|x*, X, Y)dF_Z \approx \int p(y_*|F_Z, x*, Z)q(F_Z)dF_Z \\
= \int \mathcal{N}(y|k_{xZ}{K_{ZZ}}^{-1}F_Z, k_{xx} - k_{xZ}{K_{ZZ}}^{-1}k_{Zx} + \sigma^2\mathcal{N}(F_Z|\mu, \Sigma)dF_Z \\
= \mathcal{N}(y|k_{xZ}{K_{ZZ}}^{-1}\mu, k_{xx} - k_{xZ}{K_{ZZ}}^{-1}k_{Zx} + \sigma^2 + k_{xZ}{K_{ZZ}}^{-1}\Sigma{K_{ZZ}}^{-1}k_{Zx})
\end{aligned} \tag{4}$$

となる。

なお、定義を踏まえると、${k_{xZ}}^T = k_{Zx}, {{K_{ZZ}}^{-1}}^T = {K_{ZZ}}^{-1}$。

形は異なるが、(7.67) が求まった?

($p(f_*|Y)$ と考えると一致する。)

論文 [115] の 2 章について、少し詳しく確認する。(4) 式をスカラの x,y でなく、ベクトルとして考え、各事象は独立だと考えると、以下のように (5) 式になる。

$$p(Y|X, Z, F_Z) = \prod_{n=1}^{N} p(y_n|x_n, Z, F_Z) = \mathcal{N}(Y|K_{XZ}{K_{ZZ}}^{-1}K_{ZX}F_Z, \Lambda + \sigma^2 I) \tag{5}$$

ここで、

$$\begin{aligned}
\Lambda = diag(\lambda_n) \\
\lambda_n = k_{x_n, x_n} - k_{x_n Z}{K_{ZZ}}^{-1}k_{Z x_n}
\end{aligned} \tag{6}$$

そもそも、$\{Z, F_Z\}$ に関しても、ガウス過程になっており、F_Z は Z のみに依存するので、

$$p(F_Z|Z) = p(F_Z|X, Z) = \mathcal{N}(F_Z|0, K_{ZZ}) \tag{7}$$

一般に以下が成り立つ。

$$p(F_Z|X, Y, Z) = \frac{p(F_Z, Y|X, Z)}{p(Y|X, Z)} = \frac{p(Y|X, Z, F_Z)p(F_Z|X, Z)}{p(Y|X, Z)} \tag{8}$$

F_Z を変数、それ以外を定数 C として、以下のようになる。(正規分布の共分散行列は対称行列であること、スカラになっている部分があることに注意する。)

$$
\begin{aligned}
ln\, p(F_Z|X,Y,Z) &= -\frac{1}{2}(F_Z^T {K_{ZZ}}^{-1} F_Z + (Y - K_{XZ}{K_{ZZ}}^{-1}F_Z)^T(\Lambda+\sigma^2 I)^{-1}(Y - K_{XZ}{K_{ZZ}}^{-1}F_Z)) + C \\
&= -\frac{1}{2}(F_Z^T({K_{ZZ}}^{-1} + {K_{ZZ}}^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}K_{XZ}{K_{ZZ}}^{-1})F_Z - 2F_Z{K_{ZZ}}^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y) + C \\
&= -\frac{1}{2}(F_Z^T{K_{ZZ}}^{-1}(K_{ZZ} + K_{ZX}(\Lambda+\sigma^2 I)^{-1}K_{XZ}){K_{ZZ}}^{-1}F_Z - 2F_Z{K_{ZZ}}^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y) + C \\
&= -\frac{1}{2}(F_Z - K_{ZZ}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y)^T{K_{ZZ}}^{-1}Q{K_{ZZ}}^{-1}(F_Z - K_{ZZ}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y) + C
\end{aligned} \tag{9}
$$

これは確率分布になっているため、C で適当に正規化され、(7) 式のようにガウス分布になる。

$$
p(F_Z|X,Y,Z) = \mathcal{N}(F_Z|K_{ZZ}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y, K_{ZZ}Q^{-1}K_{ZZ}) \tag{10}
$$

(8) 式に関して、近似が入っているような気がするが、

$$
\begin{aligned}
p(y_*|x_*,Y,X,Z) &= \int p(y_*|F_Z,x_*,Y,X,Z)p(F_Z|x_*,Y,X,Z)dF_Z \\
&\approx \int p(y_*|F_Z,x_*,Z)p(F_Z|x_*,Y,X,Z)dF_Z = \int p(y_*|F_Z,x_*,Z)p(F_Z|Y,X,Z)dF_Z \\
&= \int \mathcal{N}(y_*|k_{x_*Z}{K_{ZZ}}^{-1}F_Z, k_{x_*x_*} - k_{x_*Z}{K_{ZZ}}^{-1}k_{Zx_*} + \sigma^2)\mathcal{N}(F_Z|K_{ZZ}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y, K_{ZZ}Q^{-1}K_{ZZ})dF_Z \\
&= \mathcal{N}(y_*|k_{x_*Z}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y, k_{x_*x_*} - k_{x_*Z}{K_{ZZ}}^{-1}k_{Zx_*} + \sigma^2 + k_{x_*Z}{K_{ZZ}}^{-1}K_{ZZ}Q^{-1}K_{ZZ}{K_{ZZ}}^{-1}k_{Zx_*}) \\
&= \mathcal{N}(y_*|k_{x_*Z}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y, k_{x_*x_*} - k_{x_*Z}{K_{ZZ}}^{-1}k_{Zx_*} + \sigma^2 + k_{x_*Z}Q^{-1}k_{Zx_*}) \\
&= \mathcal{N}(y_*|k_{x_*Z}Q^{-1}K_{ZX}(\Lambda+\sigma^2 I)^{-1}Y, k_{x_*x_*} - k_{x_*Z}({K_{ZZ}}^{-1} - Q^{-1})k_{Zx_*} + \sigma^2)
\end{aligned} \tag{11}
$$

となり、(8) 式が求まる。