7.2.2.2 のラプラス近似を用いた予測分布の近似について詳細を確認する。この付近については PRML の 6.4、特に 6.4.6 で同様の議論されている。

一般に、

$$p(f_*|Y,X,x_*) = \int p(f_*|Y,X,x_*,F)p(F|Y,X,x_*)dF \tag{1}$$

となるが、ガウス過程の仮定より、

$$p(f_*|Y,X,x_*,F) = p(f_*|X,x_*,F) \tag{2}$$

F は X,Y のみから推定されるので、

$$p(F|Y,X,x_*) = p(F|Y,X) \tag{3}$$

これをまとめると、

$$p(f_*|Y,X,x_*) = \int p(f_*|X,x_*,F)p(F|Y,X)dF \tag{4}$$

となる。

ここで、(7.23) を用いて、

$$p(F|Y,X) \approx q(F|X,Y) = \mathcal{N}(F|\hat{F},\Lambda^{-1}) \tag{5}$$

よって、$p(f_*|Y,X,x_*)$ も近似し、(7.31) のように、

$$p(f_*|Y,X,x_*) \approx q(f_*|Y,X,x_*) \equiv \int p(f_*|X,x_*,F)q(F|X,Y)dF \tag{6}$$

となる。

ここで、(7.20) の下も考慮し、ガウス過程から $p(x_*,X|f_*,F)$ を考える。

$$p(f_*,F|x_*,X) = \mathcal{N}\left(\begin{bmatrix} f_* \\ F \end{bmatrix}|0, \begin{bmatrix} k_\beta(x_*,x_*) & {k_*}^T \\ k_* & {K_\beta}^{-1} \end{bmatrix}\right) \tag{7}$$

これは、(7.7) と同形状で、(7.8) のように

$$p(f_*|x_*,X,F) = \mathcal{N}(f_*|{k_*}^T {K_\beta}^{-1}F, k_\beta(x_*,x_*) - {k_*}^T {K_\beta}^{-1}k_*) \tag{8}$$

よって、(A.24) も考慮して、

$$q(f_*|Y,X,x_*) = \int \mathcal{N}(f_*|{k_*}^T {K_\beta}^{-1}F, k_\beta(x_*,x_*) - {k_*}^T {K_\beta}^{-1}k_*)\mathcal{N}(F|\hat{F},\Lambda^{-1})dF = \mathcal{N}(f_*|\mu_*,{\sigma_*}^2) \tag{9}$$

ここで、

$$\mu_* = {k_*}^T {K_\beta}^{-1}\hat{F} \tag{10}$$

K_β が対称行列であることも踏まえ、

$$\begin{aligned} {\sigma_*}^2 &= k_\beta(x_*,x_*) - {k_*}^T {K_\beta}^{-1}k_* + {k_*}^T {K_\beta}^{-1}\Lambda^{-1}({k_*}^T {K_\beta}^{-1})^T \\ &= k_\beta(x_*,x_*) - {k_*}^T {K_\beta}^{-1}k_* + {k_*}^T {K_\beta}^{-1}\Lambda^{-1}{K_\beta}^{-1}k_* \\ &= k_\beta(x_*,x_*) - {k_*}^T({K_\beta}^{-1} - {K_\beta}^{-1}\Lambda^{-1}{K_\beta}^{-1})k_* \end{aligned} \tag{11}$$

さて、(7.25) を見ると (7.28) は Λ を表していることがわかる。(7.30) を復習すると、これは対角行列になっており、これを-W とすると、

$$\Lambda = W + {K_\beta}^{-1} \tag{12}$$

これを踏まえると、

$$\sigma_*{}^2 = k_\beta(x_*, x_*) - k_*{}^T({K_\beta}^{-1} - {K_\beta}^{-1}(W + {K_\beta}^{-1})^{-1}{K_\beta}^{-1})k_* \tag{13}$$

この式のカッコの中を (A.1) と見比べると、$A = K_\beta, U = I, B = W^{-1}, V = I$ の対応がつくことがわかり、その結果、

$$\sigma_*{}^2 = k_\beta(x_*, x_*) - k_*{}^T(K_\beta + W^{-1})^{-1}{K_\beta}^{-1})k_* \tag{14}$$

となることがわかる。(本に誤りがあるのでは??PRML では少し、変数の割り当て方が違うが、これと同様になっている。)