A.1.3 を確認する。(A.14),(A.15) を踏まえると、(A.16) は x の関数と考えると以下のようになる。なお、共分散行列は対称行列であることに注意する。

$$\begin{aligned}
ln\,p(x|y) = ln\,\frac{p(x,y)}{p(y)} &= ln\,p(x,y) + c = ln\,p(y|x)p(x) + c \\
= -\frac{1}{2}\{(y-(Wx+b))^T{\Sigma_y}^{-1}(y-(Wx+b))\} &- \frac{1}{2}\{(x-\mu)^T{\Sigma_x}^{-1}(x-\mu)\} + c \\
= -\frac{1}{2}\{x^TW^T{\Sigma_y}^{-1}Wx + 2x^TW{\Sigma_y}^{-1}(y-b)\} &- \frac{1}{2}\{x^T{\Sigma_x}^{-1}x - 2x^T{\Sigma_x}^{-1}\mu\} + c \\
= -\frac{1}{2}\{x^T({\Sigma_x}^{-1} + W^T{\Sigma_y}^{-1}W)x &- 2x^T({\Sigma_x}^{-1}\mu + W{\Sigma_y}^{-1}(y-b))\} + c \\
= -\frac{1}{2}\{(x-({\Sigma_x}^{-1}+W^T{\Sigma_y}^{-1}W)^{-1}({\Sigma_x}^{-1}\mu &+ W{\Sigma_y}^{-1}(y-b))^T({\Sigma_x}^{-1}+W^T{\Sigma_y}^{-1}W) \\
(x-({\Sigma_x}^{-1}+W^T{\Sigma_y}^{-1}W)^{-1}({\Sigma_x}^{-1}\mu &+ W{\Sigma_y}^{-1}(y-b))\} + c
\end{aligned} \tag{1}$$

この結果から確率 $p(x|y)$ はガウス分布となり、(A.17) から (A.19) がわかる。

ベイズの定理からわかる (A.20) も利用して、(A.21) を考える。

$$ln\,p(y) = ln\,p(y|x) - ln\,p(x|y) + c \tag{2}$$

y の関数だということに注意すると、

$$ln\,p(y|x) = -\frac{1}{2}\{(y-(Wx+b))^T{\Sigma_y}^{-1}(y-(Wx+b))\} + c = -\frac{1}{2}\{y^T{\Sigma_y}^{-1}y - 2y^T{\Sigma_y}^{-1}(Wx+b)\} + c \tag{3}$$

$$\begin{aligned}
ln\,p(x|y) = -\frac{1}{2}\{(x-\mu_{x|y})^T{\Sigma_{x|y}}^{-1}(x-\mu_{x|y})\} + c &= -\frac{1}{2}\{-2{\mu_{x|y}}^T{\Sigma_{x|y}}^{-1}x + {\mu_{x|y}}^T{\Sigma_{x|y}}^{-1}\mu_{x|y}\} + c \\
= -\frac{1}{2}\{-2({\Sigma_x}^{-1}\mu + W^T{\Sigma_y}^{-1}(y-b))^Tx + ({\Sigma_x}^{-1}\mu &+ W^T{\Sigma_y}^{-1}(y-b))^T\Sigma_{x|y}({\Sigma_x}^{-1}\mu + W^T{\Sigma_y}^{-1}(y-b))\} + c \\
= -\frac{1}{2}\{y^T{\Sigma_y}^{-1}W\Sigma_{x|y}W^T{\Sigma_y}^{-1}y &- 2y^T{\Sigma_y}^{-1}W(x + \Sigma_{x|y}(W^T{\Sigma_y}^{-1}b - {\Sigma_x}^{-1}\mu))\} + c
\end{aligned} \tag{4}$$

よって、

$$\begin{aligned}
ln\,p(y) = ln\,p(y|x) - ln\,p(x|y) + c &= -\frac{1}{2}\{y^T{\Sigma_y}^{-1}y - 2y^T{\Sigma_y}^{-1}(Wx+b)\} + \\
\frac{1}{2}\{y^T{\Sigma_y}^{-1}W\Sigma_{x|y}W^T{\Sigma_y}^{-1}y &- 2y^T{\Sigma_y}^{-1}W(x + \Sigma_{x|y}(W^T{\Sigma_y}^{-1}b - {\Sigma_x}^{-1}\mu))\} + c \\
= -\frac{1}{2}\{y^T({\Sigma_y}^{-1} - {\Sigma_y}^{-1}W\Sigma_{x|y}W^T{\Sigma_y}^{-1})y &- 2y^T{\Sigma_y}^{-1}(b - W\Sigma_{x|y}(W^T{\Sigma_y}^{-1}b - {\Sigma_x}^{-1}\mu))\} + c \\
= -\frac{1}{2}\{y^T({\Sigma_y}^{-1} - {\Sigma_y}^{-1}W\Sigma_{x|y}W^T{\Sigma_y}^{-1})y &- 2y^T{\Sigma_y}^{-1}(b + W\Sigma_{x|y}({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b))\} + c
\end{aligned} \tag{5}$$

ここで (A.1) を用いると、

$$A \equiv {\Sigma_y}^{-1} - {\Sigma_y}^{-1}W\Sigma_{x|y}W^T{\Sigma_y}^{-1} = {\Sigma_y}^{-1} - {\Sigma_y}^{-1}W({\Sigma_x}^{-1} + W^T{\Sigma_y}^{-1}W)^{-1}W^T{\Sigma_y}^{-1} = (\Sigma_y + W\Sigma_xW^T)^{-1} \tag{6}$$

また、(A.2) を用いると、

$$
\begin{aligned}
B \equiv A^{-1}{\Sigma_y}^{-1}(b + W\Sigma_{x|y}({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b)) \\
= (\Sigma_y + W\Sigma_x W^T){\Sigma_y}^{-1}(b + W({\Sigma_x}^{-1} + W^T{\Sigma_y}^{-1}W)^{-1}({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b)) \\
= (\Sigma_y + W\Sigma_x W^T)({\Sigma_y}^{-1}b + {\Sigma_y}^{-1}W({\Sigma_x}^{-1} + W^T{\Sigma_y}^{-1}W)^{-1}({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b)) \\
= (\Sigma_y + W\Sigma_x W^T)({\Sigma_y}^{-1}b + (\Sigma_y + W\Sigma_x W^T)^{-1}W\Sigma_x({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b)) \\
= b + W\Sigma_x W^T{\Sigma_y}^{-1}b + W\Sigma_x({\Sigma_x}^{-1}\mu - W^T{\Sigma_y}^{-1}b) = b + W\Sigma_x W^T{\Sigma_y}^{-1}b + W\mu - W\Sigma_x W^T{\Sigma_y}^{-1}b \\
= W\mu + b
\end{aligned}
\tag{7}
$$

$$
ln\, p(y) = -\frac{1}{2}\{y^T Ay - 2y^T AB\} + c = -\frac{1}{2}\{(y-B)^T A(y-B)\} + c \tag{8}
$$

よって、$p(y)$ は (A.24) で表される。

さて、ここで $p(y_*|x_*, X, Y)$ を考えると、(3.71) が (A.14)、(3.67) が (A.15) に対応するとして、(A.21) を考える。

すると、(A.14) の x が w,μ が $\hat{\mu}$,Σ_x が $\hat{\Sigma}$ に (A.15) の W が $\phi(x_*)^T$, b が 0,Σ_y が σ^{-2}, y が y_* に対応する。

$$
\begin{aligned}
ln\, p(y_*|x_*, Y, X) = -\frac{1}{2}\{y_*({\sigma_y}^{-2} - {\sigma_y}^{-2}\phi(x_*)^T(\hat{\Sigma}^{-1} + \phi(x_*)\sigma_y^{-2}\phi(x_*)^T)^{-1}\phi(x_*)\sigma_y^{-2})y_* \\
-2y_*{\sigma_y}^{-2}\phi(x_*)^T(\hat{\Sigma}^{-1} + \phi(x_*)\sigma_y^{-2}\phi(x_*)^T)^{-1}\hat{\Sigma}^{-1}\mu\} + c \\
= -\frac{1}{2}\{({\sigma_y}^{-2} - {\sigma_y}^{-4}\phi(x_*)^T(\phi(x_*)\sigma_y^{-2}\phi(x_*)^T + \hat{\Sigma}^{-1})^{-1}\phi(x_*)){y_*}^2 \\
-2\phi(x_*)^T{\sigma_y}^{-2}(\phi(x_*)\sigma_y^{-2}\phi(x_*)^T + \hat{\Sigma}^{-1})^{-1}\hat{\Sigma}^{-1}\mu y_*\} + c
\end{aligned}
\tag{9}
$$

この付近で誤植と思われるところをまとめておく。

- (A.2) に関して、行列 P,R は正定値行列でなく正則行列であれば、逆行列を持つので、(A.2) が成立すると思われる。PRML でも (C.5) に同様の式があるが、特に条件は記載されていない。
- ((A.3) は利用していないが、)(A.3) の右辺の全体には-1 は不要と思われる。PRML で (2.76) に同様の式がある。
- (A.21) について、2 番目の等号のあと、2 項目の {} の中の 2 項目は-でなく + であると思われる。前後で、符号が一致しない。
- (A.22) について、この式での式変形は (A.2) でなく、(A.1) を利用している。
- (3.74) について、y_*^2 の係数の $(\sigma_y^{-2}\phi(x_*)\phi(x_*)^T + \hat{\Sigma}^{-1})$ は $(\sigma_y^{-2}\phi(x_*)\phi(x_*)^T + \hat{\Sigma}^{-1})^{-1}$ であると思われる。(A.21) に代入するとそうなる。