A.4.1 を確認する。

7.4.2 の (7.87),(7.88) 式にて、再帰的な定義が得られている。しかし、この式には活性化関数として、一般的な非線形関数 $\phi(a)$ が含まれているため、(7.89) で表される期待値の部分で解析的な計算ができない。

$\phi(a)$ が特定の非線形関数のときは (7.89) の部分も解析的に求められ、$\phi(a) = Erf(a)$ のときは 7.4.3.1 に、$\phi(a) = \Theta(a)a^m$ のときは、7.4.3.2 に示されている。

ここでは 7.4.3.1 に関して、詳細が記述されている A.4.1 を確認する。

7.4.3 にあるように、a_1, a_2 は入力 x, x' に対する活性であり、それぞれ、スカラとなる。それを要素とする a は 2 次元のベクトルとなり、(7.83) の上にあるようにそれは 0 を平均とする、ガウス分布と仮定しており、その分散は 2 次元正方行列の Σ で表される。当然、これは対称行列になっている。

(7.90) にあるように活性化関数を

$$\phi(a) = Erf(a) = \frac{2}{\sqrt{\pi}} \int_0^a exp(-t^2)dt \tag{1}$$

としたときを考える。

(A.45) のように、

$$\begin{aligned} C &= \mathbb{E}_{a\sim\mathcal{N}(0,\Sigma)}[\phi(a_1)\phi(a_2)] = \int \mathcal{N}(a|0,\Sigma)Erf(a_1)Erf(a_2)da = \int \mathcal{N}(a|0,\Sigma)\phi(a_1)\phi(a_2)da \\ &= \int (\frac{1}{(\sqrt{2\pi})^2\sqrt{|\Sigma|}} exp(-\frac{1}{2}a^T\Sigma^{-1}a))\phi(e_1^Ta)\phi(e_2^Ta)da = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \int exp(-\frac{1}{2}a^T\Sigma^{-1}a)\phi(e_1^Ta)\phi(e_2^Ta)da \end{aligned} \tag{2}$$

ただし、$e_1 = (1,0)^T, e_2 = (0,1)^T$。

Σ は共分散行列で、対称行列なので、コレスキー分解ができ、$\Sigma = LL^T$ とおける。なお L は下三角行列になる。変数変換で $L\hat{a} = a$ とし、$b_1^T = e_1^T L, b_2^T = e_2^T L$ として、

$$\begin{aligned} C &= \frac{1}{2\pi||L||L^T||^{\frac{1}{2}}} \int exp(-\frac{1}{2}\hat{a}^T L^T ({L^T}^{-1} L^{-1}) L\hat{a})\phi(e_1^T L\hat{a})\phi(e_2^T L\hat{a})(|L|d\hat{a}) \\ &= \frac{|L|}{2\pi||L|^2|^{\frac{1}{2}}} \int exp(-\frac{1}{2}\hat{a}^T\hat{a})\phi(e_1^T L\hat{a})\phi(e_2^T L\hat{a})d\hat{a} = \frac{1}{2\pi} \int exp(-\frac{1}{2}\hat{a}^T\hat{a})\phi(b_1^T\hat{a})\phi(b_2^T\hat{a})d\hat{a} \end{aligned} \tag{3}$$

(A.47) のように、補助変数を導入する。

$$C(\lambda) = \frac{1}{2\pi} \int exp(-\frac{1}{2}\hat{a}^T\hat{a})\phi(\lambda b_1^T\hat{a})\phi(b_2^T\hat{a})d\hat{a} \tag{4}$$

ここで、$C(1) = C$。また、

$$Erf(0) = \frac{2}{\sqrt{\pi}} \int_0^0 exp(-t^2)dt = 0 \tag{5}$$

なので、

$$C(0) = \frac{1}{2\pi} \int exp(-\frac{1}{2}\hat{a}^T\hat{a})\phi(0)\phi(b_2^T\hat{a})d\hat{a} = \frac{1}{2\pi} \int exp(-\frac{1}{2}\hat{a}^T\hat{a})0\phi(b_2^T\hat{a})d\hat{a} = 0 \tag{6}$$

(A.49) を踏まえて、$C(\lambda)$ を λ で微分すると、(A.48) のようになる。

$$
\begin{aligned}
C'(\lambda) &= \frac{1}{2\pi}\int exp(-\frac{1}{2}\hat{a}^T\hat{a})\frac{d}{d\lambda}\phi(\lambda b_1^T\hat{a})\phi(b_2^T\hat{a})d\hat{a} = \frac{1}{2\pi}\int exp(-\frac{1}{2}\hat{a}^T\hat{a})\frac{d}{d(\lambda b_1^T\hat{a})}\phi(\lambda b_1^T\hat{a})\frac{d}{d\lambda}\lambda b_1^T\hat{a}\phi(b_2^T\hat{a})d\hat{a} \\
&= \frac{1}{2\pi}\int exp(-\frac{1}{2}\hat{a}^T\hat{a})\frac{2}{\sqrt{\pi}}exp(-(\lambda b_1^T\hat{a})^2)b_1^T\hat{a}\phi(b_2^T\hat{a})d\hat{a} = \frac{1}{2\pi}\frac{2}{\sqrt{\pi}}\int exp(-\frac{1}{2}\hat{a}^T\hat{a}-(\lambda^2\hat{a}^T b_1 b_1^T\hat{a}))b_1^T\hat{a}\phi(b_2^T\hat{a})d\hat{a} \\
&= \frac{1}{\pi^{\frac{3}{2}}}\int exp(-\frac{1}{2}\hat{a}^T(I+2\lambda^2 b_1 b_1^T)\hat{a})b_1^T\hat{a}\phi(b_2^T\hat{a})d\hat{a}
\end{aligned} \tag{7}
$$

$$
(I+2\lambda^2 b_1 b_1^T)^T = I+2\lambda^2 b_1 b_1^T \tag{8}
$$

なので、$I+2\lambda^2 b_1 b_1^T$ は対称行列で、逆行列も対称行列となり、これもコレスキー分解ができる。$(I+2\lambda^2 b_1 b_1^T)^{-1} = L_b {L_b}^T$($L_b$ は下三角行列) とし、変数変換 $L_b\tilde{a} = \hat{a}$、$c_1^T = b_1^T L_b, c_2^T = b_2^T L_b$ とすると、

$$
C'(\lambda) = \frac{1}{\pi^{\frac{3}{2}}}\int exp(-\frac{1}{2}\tilde{a}^T L_b^T {L_b^T}^{-1} L_b^{-1} L_b\tilde{a})b_1^T L_b\tilde{a}\phi(b_2^T L_b\tilde{a})|L_b|d\tilde{a} = \frac{|L_b|}{\pi^{\frac{3}{2}}}\int exp(-\frac{1}{2}\tilde{a}^T\tilde{a})c_1^T\tilde{a}\phi(c_2^T\tilde{a})d\tilde{a} \tag{9}
$$

これにより、(A.50) が求まった。

(A.51) に関して、要素に分解してみる。

$$
\begin{aligned}
C'(\lambda) &= \frac{2|L_b|{c_1}^T c_2}{\pi^2}\int exp(-\frac{1}{2}\tilde{a}^T(I+2c_2{c_2}^T)\tilde{a})d\tilde{a} \\
&= \frac{2|L_b|(c_{1,1}c_{2,1}+c_{1,2}c_{2,2})}{\pi^2}\int exp(-\frac{1}{2}\tilde{a}^T\begin{pmatrix}1+2c_{2,1}^2 & 2c_{2,1}c_{2,2}\\ 2c_{2,1}c_{2,2} & 1+2c_{2,2}^2\end{pmatrix}\tilde{a})d\tilde{a} \\
&= \frac{2|L_b|(c_{1,1}c_{2,1}+c_{1,2}c_{2,2})}{\pi^2}\int\int exp(-\frac{1}{2}(\tilde{a}_1^2(1+2c_{2,1}^2)+4\tilde{a}_1\tilde{a}_2 c_{2,1}c_{2,2}+\tilde{a}_2^2(1+2c_{2,2}^2))d\tilde{a}_1 d\tilde{a}_2 \\
&= \frac{2|L_b|(c_{1,1}c_{2,1}+c_{1,2}c_{2,2})}{\pi^2}\int\int exp(-\frac{1}{2}\tilde{a}_1^2)exp(-\frac{1}{2}\tilde{a}_2^2)exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_1 d\tilde{a}_2
\end{aligned} \tag{10}
$$

これが、(A.50) と等しいことを確認する。

ここで、(2.28) を振り返る。$y=\sqrt{2}t$ の変数変換を考えると、

$$
\begin{aligned}
erf(\frac{2}{\sqrt{2}}) &= \frac{2}{\sqrt{\pi}}\int_0^{\frac{x}{\sqrt{2}}} exp(-t^2)dt = \frac{2}{\sqrt{2\pi}}\int_0^x exp(-\frac{y^2}{2})dy \\
&= \frac{2}{\sqrt{2\pi}}(\int_{-\infty}^x exp(-\frac{y^2}{2})dy - \int_{-\infty}^0 exp(-\frac{y^2}{2})dy) = 2(\int_{-\infty}^x \frac{1}{\sqrt{2\pi}}exp(-\frac{y^2}{2})dy - \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}}exp(-\frac{y^2}{2})dy) \\
&= 2(\int_{-\infty}^x \mathcal{N}(y|0,1)dy - \int_{-\infty}^0 \mathcal{N}(y|0,1)dy) = 2(\Phi(x)-\Phi(0)) = 2(\Phi(x)-\frac{1}{2})
\end{aligned} \tag{11}
$$

よって、(2.28) にあるように、

$$
\Phi(x) = \frac{1}{2}(1+erf(\frac{x}{\sqrt{2}})) \tag{12}
$$

$0<\Phi(x)<1$ を考慮すると、$-\frac{1}{2}<erf(x)<\frac{1}{2}$ となる。つまり、$erf(x)$ は有限の値を取る。

(A.50) から変形する。

$$
\begin{aligned}
C'(\lambda) &= \frac{|L_b|}{\pi^{\frac{3}{2}}}\int exp(-\frac{1}{2}\tilde{a}^T\tilde{a})c_1^T\tilde{a}\phi(c_2^T\tilde{a})d\tilde{a} \\
&= \frac{|L_b|}{\pi^{\frac{3}{2}}}\int\int exp(-\frac{1}{2}(\tilde{a}_1\tilde{a}_1+\tilde{a}_2\tilde{a}_2))(c_{1,1}\tilde{a}_1+c_{1,2}\tilde{a}_2)\frac{2}{\sqrt{\pi}}\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dtd\tilde{a}_1d\tilde{a}_2 \\
&= \frac{2|L_b|}{\pi^2}(\int exp(-\frac{1}{2}\tilde{a}_2^2)\int c_{1,1}\tilde{a}_1 exp(-\frac{1}{2}\tilde{a}_1^2)\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dtd\tilde{a}_1d\tilde{a}_2 \\
&\quad +\int exp(-\frac{1}{2}\tilde{a}_1^2)\int c_{1,2}\tilde{a}_2 exp(-\frac{1}{2}\tilde{a}_2^2)\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dtd\tilde{a}_2d\tilde{a}_1) \\
&= \frac{2|L_b|}{\pi^2}(\int exp(-\frac{1}{2}\tilde{a}_2^2)c_{1,1}([-exp(-\frac{1}{2}\tilde{a}_1^2)\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dt]_{-\infty}^{\infty}+\int exp(-\frac{1}{2}\tilde{a}_1^2)c_{2,1} \\
&\quad exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_1))d\tilde{a}_2+\int exp(-\frac{1}{2}\tilde{a}_1^2)c_{1,2}([-exp(-\frac{1}{2}\tilde{a}_2^2)\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dt]_{-\infty}^{\infty} \\
&\quad +\int exp(-\frac{1}{2}\tilde{a}_2^2)c_{2,2}exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_2)d\tilde{a}_1) \\
&= \frac{2|L_b|}{\pi^2}(c_{1,1}c_{2,1}\int exp(-\frac{1}{2}\tilde{a}_2^2)\int exp(-\frac{1}{2}\tilde{a}_1^2)exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_1d\tilde{a}_2 \\
&\quad +c_{1,2}c_{2,2}\int exp(-\frac{1}{2}\tilde{a}_1^2)\int exp(-\frac{1}{2}\tilde{a}_2^2)exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_2d\tilde{a}_1) \\
&= \frac{2|L_b|(c_{1,1}c_{2,1}+c_{1,2}c_{2,2})}{\pi^2}\int\int exp(-\frac{1}{2}\tilde{a}_1^2)exp(-\frac{1}{2}\tilde{a}_2^2)exp(-(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)^2)d\tilde{a}_1d\tilde{a}_2
\end{aligned} \tag{13}
$$

となり、一致することがわかった。

$[-exp(-\frac{1}{2}\tilde{a}_2^2)\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dt]_{-\infty}^{\infty}$ に関して、積分部分は先に示したように有限の値になるので、0 になることがわかる。また、$\int_0^{c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2} exp(-t^2)dt$ の微分に関して、$(\Phi(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)-\Phi(0))' = \Phi(c_{2,1}\tilde{a}_1+c_{2,2}\tilde{a}_2)'$。

(A.51) に関して、積分部分がガウス分布の正規化定数になること、a などが 2 次元だったことを考慮すると以下のようになる。

$$
\begin{aligned}
C'(\lambda) &= \frac{2|L_b|{c_1}^Tc_2}{\pi^2}\int exp(-\frac{1}{2}\tilde{a}^T(I+2c_2{c_2}^T)\tilde{a})d\tilde{a} = \frac{2|L_b|{c_1}^Tc_2}{\pi^2}\sqrt{(2\pi)^2|I+2c_2{c_2}^T|^{-1}} \\
&= \frac{4|L_b|{c_1}^Tc_2}{\pi|I+2c_2{c_2}^T|^{\frac{1}{2}}} = \frac{4{c_1}^Tc_2}{\pi|{L_b^T}^{-1}|^{\frac{1}{2}}|I+2c_2{c_2}^T|^{\frac{1}{2}}|{L_b}^{-1}|^{\frac{1}{2}}} = \frac{4{b_1}^TL_bL_b^Tb_2}{\pi|{L_b^T}^{-1}{L_b}^{-1}+2{L_b^T}^{-1}L_b^Tb_2{b_2}^TL_b{L_b}^{-1}|^{\frac{1}{2}}} \\
&= \frac{4{b_1}^TL_bL_b^Tb_2}{\pi|(L_bL_b^T)^{-1}+2b_2{b_2}^T|^{\frac{1}{2}}} = \frac{4{b_1}^TL_bL_b^Tb_2}{\pi|I+2\lambda^2b_1b_1^T+2b_2{b_2}^T|^{\frac{1}{2}}} = \frac{4{b_1}^TL_bL_b^Tb_2}{\pi\Delta^{\frac{1}{2}}}
\end{aligned} \tag{14}
$$

ここで

$$
\begin{aligned}
\Delta &= |I+2\lambda^2b_1b_1^T+2b_2{b_2}^T| = \begin{vmatrix} 1+2\lambda^2b_{1,1}^2+2b_{2,1}^2 & 2\lambda^2b_{1,1}b_{1,2}+2b_{2,1}b_{2,2} \\ 2\lambda^2b_{1,1}b_{1,2}+2b_{2,1}b_{2,2} & 1+2\lambda^2b_{1,2}^2+2b_{2,2}^2 \end{vmatrix} \\
&= 1+2\lambda(b_{1,1}^2+b_{1,2}^2)+2(b_{2,1}^2+b_{2,2}^2)+4\lambda^2b_{2,1}^2b_{1,2}^2+4\lambda^2b_{1,1}^2b_{2,2}^2-8\lambda^2b_{1,1}b_{1,2}b_{2,1}b_{2,2} \\
&= 1+2\lambda^2b_1^Tb_1+2b_2^Tb_2+4\lambda^2(b_1^Tb_1b_2^Tb_2-(b_1^Tb_2)^2)
\end{aligned} \tag{15}
$$

となる。((A.55) は違っているのでは?)

シャーマン・モリソンの公式も確認しておく。

$$\begin{aligned} I = (I+gg^T)(I+gg^T)^{-1} = (I+gg^T)(I-\frac{gg^T}{1+g^Tg}) \\ = I+gg^T-\frac{gg^T}{1+g^Tg}-\frac{gg^Tgg^T}{1+g^Tg} = Igg^T-\frac{gg^T+g(g^Tg)g^T}{1+g^Tg} = Igg^T-\frac{(1+g^Tg)gg^T}{1+g^Tg} \\ = Igg^T-gg^T = I \end{aligned} \tag{16}$$

この式を用いると (A.54) が言える。

$$\begin{aligned} b_1^TL_bL_b^Tb_2 = b_1^T(I+2\lambda b_1b_1^T)^{-1}b_2 = b_1^T(I-\frac{2\lambda^2b_1b_1^T}{1+2\lambda^2b_1^Tb_1})b_2 = b_1^Tb_2-\frac{2\lambda^2b_1^Tb_1b_1^Tb_2}{1+2\lambda^2b_1^Tb_1} \\ = \frac{b_1^Tb_2+2\lambda^2b_1^Tb_1b_1^Tb_2-2\lambda^2b_1^Tb_1b_1^Tb_2}{1+2\lambda^2b_1^Tb_1} = \frac{b_1^Tb_2}{1+2\lambda^2b_1^Tb_1} \end{aligned} \tag{17}$$

これを踏まえると、

$$C'(\lambda) = \frac{4{b_1}^TL_bL_b^Tb_2}{\pi\Delta^{\frac{1}{2}}} = \frac{4\frac{b_1^Tb_2}{1+2\lambda^2b_1^Tb_1}}{\pi\Delta^{\frac{1}{2}}} = \frac{4b_1^Tb_2}{\pi(1+2\lambda^2b_1^Tb_1)\Delta^{\frac{1}{2}}} \tag{18}$$

となり、(A.56) は成立する。

(A.57) のように z を以下のように置く。

$$z = \frac{2\lambda b_1^Tb_2}{\sqrt{(1+2\lambda^2b_1^Tb_1)(1+2b_2^Tb_2)}} \tag{19}$$

このとき、

$$\begin{aligned} \frac{dz}{d\lambda} = \frac{2\lambda b_1^Tb_2}{\sqrt{(1+2\lambda^2b_1^Tb_1)(1+2b_2^Tb_2)}} = (\frac{2b_1^Tb_2}{\sqrt{1+2b_2^Tb_2}}\frac{\lambda}{\sqrt{1+2\lambda^2b_1^Tb_1}})' \\ = \frac{2b_1^Tb_2}{\sqrt{1+2b_2^Tb_2}}(\frac{1}{\sqrt{1+2\lambda^2b_1^Tb_1}}-\frac{2\lambda b_1^Tb_1}{\sqrt{1+2\lambda^2b_1^Tb_1}^{\frac{3}{2}}}) = \frac{2b_1^Tb_2}{\sqrt{1+2b_2^Tb_2}}(\frac{1}{\sqrt{1+2\lambda^2b_1^Tb_1}^{\frac{3}{2}}}) \\ = \frac{2b_1^Tb_2}{\sqrt{1+2b_2^Tb_2}\sqrt{1+2\lambda^2b_1^Tb_1}^{\frac{3}{2}}} \end{aligned} \tag{20}$$

このとき、(A.58) が成り立つことを以下で確認する。

$$\begin{aligned} C'(\lambda) = \frac{2}{\pi\sqrt{1-z^2}}\frac{dz}{d\lambda} = \frac{2}{\pi\sqrt{1-\frac{4\lambda^2(b_1^Tb_2)^2}{(1+2b_2^Tb_2)(1+2\lambda b_1^Tb_1)}}}\frac{dz}{d\lambda} \\ = \frac{2\sqrt{(1+2b_2^Tb_2)(1+2\lambda b_1^Tb_1)}}{\pi\sqrt{1+2\lambda^2b_1^Tb_1+2b_2^Tb_2+4\lambda^2(b_1^Tb_1b_2^Tb_2-(b_1^Tb_2)^2)}}\frac{2b_1^Tb_2}{\sqrt{1+2b_2^Tb_2}\sqrt{1+2\lambda^2b_1^Tb_1}^{\frac{3}{2}}} \\ = \frac{4b_1^Tb_2}{\pi\sqrt{1+2\lambda^2b_1^Tb_1+2b_2^Tb_2+4\lambda^2(b_1^Tb_1b_2^Tb_2-(b_1^Tb_2)^2)}(1+2\lambda^2b_1^Tb_1)} = \frac{4b_1^Tb_2}{\pi\Delta^{\frac{1}{2}}(1+2\lambda^2b_1^Tb_1)} \end{aligned} \tag{21}$$

$y = arcsinz$ の微分を考える。$z = siny$ なので、

$$\frac{dz}{dy} = cosz = \sqrt{1-z^2} \tag{22}$$

よって、

$$\frac{dy}{dz} = \frac{d}{dz} arcsinz = \frac{1}{\sqrt{1-z^2}} \tag{23}$$

$C'(\lambda)$ を積分して、

$$C(\lambda) = \frac{2}{\pi} arcsinz + C \tag{24}$$

しかし、先に示したように $C(0) = 0$ なので、積分定数 C = 0。よって、(A.59) が成り立つことがわかる。

これを考慮すると、以下のように (A.60) が成り立つことがわかる。

$$\begin{aligned} C = C(1) &= \frac{2}{\pi} arcsin(\frac{2b_1^T b_2}{\sqrt{(1+2b_1^T b_1)(1+2b_2^T b_2)}}) = \frac{2}{\pi} arcsin(\frac{2e_1^T LL^T e_2}{\sqrt{(1+2e_1^T LL^T e_1)(1+2e_2^T LL^T e_2)}}) \\ &= \frac{2}{\pi} arcsin(\frac{2e_1^T \Sigma e_2}{\sqrt{(1+2e_1^T \Sigma e_1)(1+2e_2^T \Sigma e_2)}}) = \frac{2}{\pi} arcsin(\frac{2\Sigma_{1,2}}{\sqrt{(1+2\Sigma_{1,1})(1+2\Sigma_{2,2})}}) \end{aligned} \tag{25}$$