機械学習
正誤対応ならびに追記とコメント

共立出版

最終更新：2024 年 4 月 16 日

正誤表

巻	ページ	行数, 図・表・式番号	誤	正
1	11	1 行目	$\|\mathbf{w}\|^2 = \sqrt{w_0^2 + w_1^2 + w_2^2 + w_3^2}$	$\|\mathbf{w}\|^2 = w_0^2 + w_1^2 + w_2^2 + w_3^2$
1	14	第 3 パラグラフ, 3 行目から 4 行目	（追記）	下の追記と注釈の表を参照
1	23	1.3.3 確率の解釈：頻度主義とベイズ主義	（追記）	下の追記と注釈の表を参照
1	29	2 行目	（追記）	下の追記と注釈の表を参照
1	30	5 行目 式	$\cdots = \operatorname*{argmax}_{\mathbf{w}} p(\mathcal{D} \mid \mathbf{w}) p(w)$	$\cdots = \operatorname*{argmax}_{\mathbf{w}} p(\mathcal{D} \mid \mathbf{w}) p(\mathbf{w})$
1	32	下から 8 行目から 7 行目	（追記）	下の追記と注釈の表を参照
1	34	下から 4 行目（脚注を除く）	（追記）	下の追記と注釈の表を参照
1	58	脚注 下から 2 行目	勾配 $E(\mathbf{w})$ は	勾配 $\nabla E(\mathbf{w})$ は
1	60	2 行目	ベクトル $\mathbf{x}$ による微分	ベクトル $\mathbf{x}^{\mathrm{T}}$ による微分
1	60	7 行目 式 (2.2.6)	$\dfrac{d}{d\mathbf{x}} \mathbf{A}\mathbf{x} = \mathbf{A}$	$\dfrac{d}{d\mathbf{x}^{\mathrm{T}}} \mathbf{A}\mathbf{x} = \mathbf{A}$
1	60	10 行目 式 (2.2.7)	$\dfrac{d\mathbf{x}}{d\mathbf{x}} = \dfrac{d}{d\mathbf{x}} \mathbf{I}\mathbf{x} = \mathbf{I}$	$\dfrac{d\mathbf{x}}{d\mathbf{x}^{\mathrm{T}}} = \dfrac{d}{d\mathbf{x}^{\mathrm{T}}} \mathbf{I}\mathbf{x} = \mathbf{I}$
1	60	15 行目 式 (2.2.8)	$\dfrac{df(\boldsymbol{\mu})}{d\boldsymbol{\mu}} = \dfrac{df(\mathbf{m})}{d\mathbf{m}} \dfrac{d\mathbf{m}}{d\boldsymbol{\mu}}$ $= \dfrac{d}{d\mathbf{m}} \mathbf{m}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{m} \cdot \dfrac{d}{d\boldsymbol{\mu}} (\mathbf{x}_n - \boldsymbol{\mu})$ $= 2\boldsymbol{\Sigma}^{-1} \mathbf{m} \cdot (-\mathbf{I}) = -2\boldsymbol{\Sigma}^{-1} (\mathbf{x}_n - \boldsymbol{\mu})$	$\dfrac{df(\boldsymbol{\mu})}{d\boldsymbol{\mu}} = \dfrac{d\mathbf{m}^{\mathrm{T}}}{d\boldsymbol{\mu}} \dfrac{df(\mathbf{m})}{d\mathbf{m}}$ $= \dfrac{d}{d\boldsymbol{\mu}} (\mathbf{x}_n - \boldsymbol{\mu})^{\mathrm{T}} \cdot \dfrac{d}{d\mathbf{m}} \mathbf{m}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{m}$ $= -\mathbf{I} \cdot 2\boldsymbol{\Sigma}^{-1} \mathbf{m} = -2\boldsymbol{\Sigma}^{-1} (\mathbf{x}_n - \boldsymbol{\mu})$
1	69	10 行目 式	$= \lambda^{N/2} \left(-\dfrac{\lambda \mu^2 N}{2} + \cdots \right)$	$= \lambda^{N/2} \exp\left(-\dfrac{\lambda \mu^2 N}{2} + \cdots \right)$
1	82	4 行目 式 (3.2.2)	$\cdots = \mathcal{N}(t \mid \mathbf{w}^{\mathrm{T}} \mathbf{x}, \sigma^2)$	$\cdots = \mathcal{N}(t \mid \boldsymbol{\phi}(\mathbf{w}^{\mathrm{T}} \mathbf{x}), \sigma^2)$

正誤表

1	85	下から 4 行目	未知数（$\mathbf{x}$ の各成分）が方程式よりも多い.	方程式が，未知数（$\mathbf{x}$ の各成分）よりも多い.
1	89	下から 2 行目 式 (3.4.4)	$\cdots = \frac{1}{(2\pi)^{\frac{M}{2}}\lvert\mathbf{S}_0\rvert}\exp\cdots$	$\cdots = \frac{1}{(2\pi)^{\frac{M}{2}}\lvert\mathbf{S}_0\rvert^{\frac{1}{2}}}\exp\cdots$
1	90	3 行目 式 (3.4.5)	$\cdots = \frac{\alpha^M}{(2\pi)^{\frac{M}{2}}}\exp\left(-\frac{\alpha\mathbf{w}^{\mathrm{T}}\mathbf{w}}{2}\right)$	$\cdots = \frac{\alpha^{\frac{M}{2}}}{(2\pi)^{\frac{M}{2}}}\exp\left(-\frac{\alpha\mathbf{w}^{\mathrm{T}}\mathbf{w}}{2}\right)$
1	90	9 行目 式 (3.4.6)	$\cdots = \frac{1}{(2\pi)^{\frac{M}{2}}\lvert\mathbf{S}_N\rvert}\exp\cdots$	$\cdots = \frac{1}{(2\pi)^{\frac{M}{2}}\lvert\mathbf{S}_N\rvert^{\frac{1}{2}}}\exp\cdots$
1	92	下から 1 行目	（追記）	下の追記と注釈の表を参照
1	101	3 行目 式	$\cdots = \iiint p(t,\,\mathbf{w},\,\alpha\,\vert\,\mathbf{x},\,\mathbf{t},\,\mathbf{X},\,\beta)\,d\mathbf{w}d\alpha$	$\cdots = \iint p(t,\,\mathbf{w},\,\alpha\,\vert\,\mathbf{x},\,\mathbf{t},\,\mathbf{X},\,\beta)\,d\mathbf{w}d\alpha$
1	102	12 行目 式 (3.4.5)	$\cdots = \frac{\alpha^M}{(2\pi)^{\frac{M}{2}}}\exp\left(-\frac{\alpha\mathbf{w}^{\mathrm{T}}\mathbf{w}}{2}\right)$	$\cdots = \frac{\alpha^{\frac{M}{2}}}{(2\pi)^{\frac{M}{2}}}\exp\left(-\frac{\alpha\mathbf{w}^{\mathrm{T}}\mathbf{w}}{2}\right)$
1	114	下から 2 行目 式	$\phi_1(x_1,\,x_2) = \exp\left\{-\frac{(x_1+1)^2+(x_2+1)^2}{2}\right\},$ $\phi_2(x_1,\,x_2) = \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\}.$	$\phi_1(x_1,\,x_2) = \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\},$ $\phi_2(x_1,\,x_2) = \exp\left\{-\frac{(x_1-1)^2+(x_2+1)^2}{2}\right\}.$
1	115	図 4.3 キャプション 4 行目	$\phi_1(\mathbf{x}) = \exp\left\{-\frac{\lvert\mathbf{x}\rvert^2}{2s^2}\right\}$	$\phi_1(\mathbf{x}) = \exp\left\{-\frac{\lVert\mathbf{x}\rVert^2}{2s^2}\right\}$
1	115	図 4.3 キャプション 5 行目	$\phi_2(\mathbf{x}) = \exp\left\{-\frac{\lvert\mathbf{x}-\boldsymbol{\mu}\rvert^2}{2s^2}\right\}$	$\phi_2(\mathbf{x}) = \exp\left\{-\frac{\lVert\mathbf{x}-\boldsymbol{\mu}\rVert^2}{2s^2}\right\}$
1	121	9 行目	（追記）	下の追記と注釈の表を参照
1	127	下から 2 行目	$\mathbf{H} = \nabla\nabla f(\mathbf{x})$	$\mathbf{H} \equiv \nabla\nabla f(\mathbf{x}) \equiv \frac{d^2 f(\mathbf{x})}{d\mathbf{x}d\mathbf{x}^{\mathrm{T}}}$
1	128	1 行目 式	$\nabla\nabla f(\mathbf{x}) = \frac{d^2 f(\mathbf{x})}{d\mathbf{x}^2} = \cdots$	$\nabla(\nabla f(\mathbf{x}))^{\mathrm{T}} = \frac{d^2 f(\mathbf{x})}{d\mathbf{x}d\mathbf{x}^{\mathrm{T}}} = \cdots$
1	133	6 行目	すなわち，尤度は無限大の「最大値」をとる.	すなわち，尤度は 1 の「最大値」をとる.
1	133	10 行目	$\sigma(k\mathbf{w}^{\mathrm{T}}\boldsymbol{\Phi}(\mathbf{x})) \to -1$ である.	$\sigma(k\mathbf{w}^{\mathrm{T}}\boldsymbol{\Phi}(\mathbf{x})) \to 0$ である.
1	136	下から 2 行目	決定境界は $x = 0.5$ であり（図 4.9 参照）,	決定境界は $w_1 x + w_0 = 0$（図 4.9 参照；図では $w_0 = 0$）,
1	138	下から 4 行目	$\cdots,\, n = 0,\,\ldots,\,N,$	$\cdots,\, n = 1,\,\ldots,\,N,$
1	147-149	付記	（注釈）	下の追記と注釈の表を参照
1	154	図 5.2(a) 左下のユニット	x_{w^2}	x_D

正誤表

1	155	12 行目 式	$\ldots, j = 1, \ldots, K$	$\ldots, k = 1, \ldots, K$
1	155	14 行目 式	$\ldots, j = 1, \ldots, K$	$\ldots, k = 1, \ldots, K$
1	167	14 行目	w_{ji} 以外の重みを固定すれば u_j で E の値が定まる．逆に，w_{ji} 以外の重みを固定すれば，E の値から u_j が定まる．	入力 $\mathbf{x}$ について，w_{ji} 以外の重みを固定したとき，w_{ji} の値に対し u_j は一意に定まり，逆に，u_j の値に対し w_{ji} は一意に定まる．
1	188	9 行目	中間層のユニットの出力を，	中間層のユニットの入力を，
1	188	下から 3 行目 式	$\tilde{u}_j^{(l)} = r_j^{(l)}\hat{z}_j^{(l)} + \beta_j^{(l)}$	$\tilde{u}_j^{(l)} = r_j^{(l)}\hat{u}_j^{(l)} + \beta_j^{(l)}$
1	199	1 行目 式	$\dfrac{d\mathbf{a}}{d\mathbf{b}} \equiv \cdots$	$\dfrac{d\mathbf{a}}{d\mathbf{b}^{\mathrm{T}}} \equiv \cdots$
1	199	2 行目	（追記）	下の追記と注釈の表を参照
1	199	4 行目 式 C1)	$\dfrac{d}{d\mathbf{x}}\mathbf{A}\mathbf{x} = \mathbf{A}.$	$\dfrac{d}{d\mathbf{x}^{\mathrm{T}}}\mathbf{A}\mathbf{x} = \mathbf{A}.$
1	199	4 行目	（追記）	下の追記と注釈の表を参照
1	199	5 行目 式 C2)	$\dfrac{d\mathbf{x}}{d\mathbf{x}} = \mathbf{I}.$	$\dfrac{d\mathbf{x}}{d\mathbf{x}^{\mathrm{T}}} = \mathbf{I}.$
1	199	5 行目	（追記）	下の追記と注釈の表を参照
1	199	下から 2 行目	ベクトル $\nabla f(\mathbf{x})$ を $\cdots$	ベクトル $(\nabla f(\mathbf{x}))^{\mathrm{T}}$ を $\cdots$
1	199	下から 1 行目 式	$\dfrac{d^2 f(\mathbf{x})}{d\mathbf{x}^2} = \nabla\nabla f(\mathbf{x}) = \cdots$	$\dfrac{d^2 f(\mathbf{x})}{d\mathbf{x}d\mathbf{x}^{\mathrm{T}}} = \nabla(\nabla f(\mathbf{x}))^{\mathrm{T}} = \cdots$
1	200	1 行目	（追記）	下の追記と注釈の表を参照
2	217	3 行目	$\lvert\mathbf{x}\rvert \cdot \lvert\mathbf{x}'\rvert \cos\theta$	$\lVert\mathbf{x}\rVert \cdot \lVert\mathbf{x}'\rVert \cos\theta$
2	220	4 行目	$r = \lVert\mathbf{x} - \mathbf{x}'\rVert$ の関数として連続ではあるが微分不可能なので，	$\mathbf{x}$ の関数として連続ではあるが， $\mathbf{x} = \mathbf{x}'$ で微分不可能なので，
2	226	脚注 下から 2 行目	$S_\omega = \{(t, X_t(\omega) \mid t \in T)\}$	$S_\omega = \{(t, X_t(\omega)) \mid t \in T\}$
2	241	4 行目	$= -\lambda g(\mathbf{x})$ が成立する必要がある．	$= -\lambda\nabla g(\mathbf{x})$ が成立する必要がある．
2	254	6 行目	(a) $\xi_n \leq 0$ のときは $\cdots$	(a) $\xi_n \leq 1$ のときは $\cdots$
2	254	7 行目	(b) $\xi_n > 0$ のときには $\cdots$	(b) $\xi_n > 1$ のときには $\cdots$
2	258	下から 8 行目	z_1 を Δz_1 に置きかえ，z_2 を $-\Delta z_1$ に置きかえて，$\cdots$	z_1 を $z_1 + \Delta z_1$ に置きかえ，z_2 を $z_2 - \Delta z_1$ に置きかえて，$\cdots$
2	259	2 行目 式 (7.4.48)	$\Delta z_1 = \begin{cases} L, & L > \frac{\beta}{\alpha}, \\ U, & U < \frac{\beta}{\alpha}, \\ \dfrac{\alpha}{\beta}, & \text{otherwise} \end{cases}$	$\Delta z_1 = \begin{cases} L, & L > \frac{\beta}{\alpha}, \\ U, & U < \frac{\beta}{\alpha}, \\ \dfrac{\beta}{\alpha}, & \text{otherwise} \end{cases}$
2	291	11 行目	この行列 $\mathbf{W}$ は，$\cdots$ 直交行列である．	一文削除

正誤表

2	292	11 行目	したがって，$\mathbf{\Sigma} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}\mathbf{x}^{\mathrm{T}}$ であり，$\cdots$	したがって，$\mathbf{\Sigma} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}$ であり，$\cdots$
2	292	下から 9 行目	（追記）	下の追記と注釈の表を参照
2	300	本文 下から 9 行目	$p \cdot (\ln p - \ln q)$ $+(1-p)\cdot(\ln(1-p) + \ln(1-q))$	$p \cdot (\ln p - \ln q)$ $+(1-p)\cdot(\ln(1-p) - \ln(1-q))$
2	329	5 行目	$\cdots = \gamma(z_{nk})^{\mathrm{old}} \cdot \left(\frac{1}{\pi_k} - \lambda\right) = 0.$	$\cdots = \gamma(z_{nk})^{\mathrm{old}} \cdot \frac{1}{\pi_k} - \lambda = 0.$
2	332	下から 6 行目	$h(\mathbf{x}) = 1$	$h(\mathbf{x}) =$ 定数
2	344	下から 8 行目 式	$\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})] \equiv \int p(\mathbf{x})f(\mathbf{x})\,d\mathbf{x} \approx \sum_{m=1}^{M} f(\hat{\mathbf{x}}_m).$	$\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})] \equiv \int p(\mathbf{x})f(\mathbf{x})\,d\mathbf{x} \approx \frac{1}{M}\sum_{m=1}^{M} f(\hat{\mathbf{x}}_m).$
3	384	4 行目 式	$J = \begin{bmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial Z} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial Z} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = -1$	$J = \begin{bmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial Z} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial Z} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = 1$
3	401	9 行目	$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = 1$	$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$
3	410	下から 3 行目	$p(\mathbf{x}) = p(\mathbf{x}_a \mid \mathbf{x}_b)$ がガウス分布であることと，$\cdots$	$p(\mathbf{x}) = p(\mathbf{x}_a, \mathbf{x}_b)$ がガウス分布であることと，$\cdots$
3	411	9 行目	ガウス分布 $p(\mathbf{x}) = p(\mathbf{x}_a \mid \mathbf{x}_b)$ の指数部を$\cdots$	ガウス分布 $p(\mathbf{x}) = p(\mathbf{x}_a, \mathbf{x}_b)$ の指数部を$\cdots$
3	420	下から 2 行目 式	$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{\Lambda}^{-1} & \mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A}\mathbf{\Lambda}^{-1} & \mathbf{L}^{-1}\mathbf{A}\mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \end{pmatrix}$	$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{\Lambda}^{-1} & \mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A}\mathbf{\Lambda}^{-1} & \mathbf{L}^{-1} + \mathbf{A}\mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \end{pmatrix}$
3	421	8 行目 式	$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{\Lambda}^{-1} & \mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A}\mathbf{\Lambda}^{-1} & \mathbf{L}^{-1}\mathbf{A}\mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \end{pmatrix}$	$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{\Lambda}^{-1} & \mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A}\mathbf{\Lambda}^{-1} & \mathbf{L}^{-1} + \mathbf{A}\mathbf{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \end{pmatrix}$
3	437	2 行目	$\ldots$, $\mathbf{A}$ を正方形行列として，$\ldots$	$\ldots$, $\mathbf{A}$ を正方行列として，$\ldots$
3	437	5 行目	$f(\mathbf{A}) = f(a_{11}, a_{12}, \ldots, a_{mn})$ で，$\cdots$	$f(\mathbf{A}) = f(a_{11}, a_{12}, \ldots, a_{nm})$ で，$\cdots$
3	445	6 行目 式	$\frac{d\mathbf{a}}{d\mathbf{b}} \equiv \cdots$	$\frac{d\mathbf{a}}{d\mathbf{b}^{\mathrm{T}}} \equiv \cdots$
3	445	7 行目	（追記）	下の追記と注釈の表を参照
3	445	9 行目 式 C1)	$\frac{d}{d\mathbf{x}}\mathbf{A}\mathbf{x} = \mathbf{A}.$	$\frac{d}{d\mathbf{x}^{\mathrm{T}}}\mathbf{A}\mathbf{x} = \mathbf{A}.$
3	445	9 行目	（追記）	下の追記と注釈の表を参照
3	445	10 行目 式 C2)	$\frac{d\mathbf{x}}{d\mathbf{x}} = \mathbf{I}.$	$\frac{d\mathbf{x}}{d\mathbf{x}^{\mathrm{T}}} = \mathbf{I}.$
3	445	10 行目	（追記）	下の追記と注釈の表を参照

正誤表

3	445	下から 1 行目	ベクトル $\nabla f(\mathbf{x})$ を $\cdots$	ベクトル $(\nabla f(\mathbf{x}))^{\mathrm{T}}$ を $\cdots$
3	446	1 行目 式	$\dfrac{d^2 f(\mathbf{x})}{d\mathbf{x}^2} = \nabla\nabla f(\mathbf{x}) = \cdots$	$\dfrac{d^2 f(\mathbf{x})}{d\mathbf{x}d\mathbf{x}^{\mathrm{T}}} = \nabla(\nabla f(\mathbf{x}))^{\mathrm{T}} = \cdots$
3	446	2 行目	（追記）	下の追記と注釈の表を参照
3	455	1 行目 式	$w = \dfrac{\sum_{n=1}^{N} t_n x_n}{\sum_{n=1}^{N} x_N^2}.$	$w = \dfrac{\sum_{n=1}^{N} t_n x_n}{\sum_{n=1}^{N} x_n^2}.$
3	456	13 行目 式	$p(\boldsymbol{\theta} \mid \mathcal{D}) \propto p(\mathcal{D} \mid w)p(w) = \cdots$	$p(\boldsymbol{\theta} \mid \mathcal{D}) \propto p(\mathcal{D} \mid \theta)p(\theta) = \cdots$
3	458	4 行目 式	$\cdots \exp\left\{-\dfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\},$	$\cdots \exp\left\{-\dfrac{1}{2}(\mathbf{x}_n-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n-\boldsymbol{\mu})\right\},$
3	458	6 行目 式	$\cdots -\dfrac{ND}{2} - \dfrac{N}{2}\ln\lvert\boldsymbol{\Sigma}\rvert.$	$\cdots -\dfrac{ND}{2}\ln(2\pi) - \dfrac{N}{2}\ln\lvert\boldsymbol{\Sigma}\rvert.$
3	458	14 行目 式	$\dfrac{\partial}{\partial\boldsymbol{\mu}}\ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \cdots$	$-\dfrac{\partial}{\partial\boldsymbol{\mu}}\ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \cdots$
3	460	2 行目 式	$= \dfrac{1}{\sqrt{2\pi}\sigma} \cdot \dfrac{1}{\sqrt{2\pi}\sigma_0} \cdots$	$= \dfrac{1}{(\sqrt{2\pi}\sigma)^N} \dfrac{1}{\sqrt{2\pi}\sigma_0} \cdots$
3	460	8-9 行目	$-\dfrac{1}{2}\left(\dfrac{N}{\sigma^2}+\dfrac{1}{\sigma_0^2}\right)$ $\left(\mu - \dfrac{\sigma_0\sum_{n=1}^{N} x_n + \mu_0\sigma}{N\sigma_0^2+\sigma^2}\right)^2 + \text{const.}$ $= -\dfrac{1}{2}\dfrac{\left(\mu - \frac{\sigma_0\sum_{n=1}^{N} x_n+\mu_0\sigma}{N\sigma_0^N+\sigma^2}\right)^2}{\frac{1}{\left(\frac{N}{\sigma^2}+\frac{1}{\sigma_0^2}\right)}} + \text{const.}$	$-\dfrac{1}{2}\left(\dfrac{N}{\sigma^2}+\dfrac{1}{\sigma_0^2}\right)$ $\left(\mu - \dfrac{\sigma_0^2\sum_{n=1}^{N} x_n + \mu_0\sigma^2}{N\sigma_0^2+\sigma^2}\right)^2 + \text{const.}$ $= -\dfrac{1}{2}\dfrac{\left(\mu - \frac{\sigma_0^2\sum_{n=1}^{N} x_n+\mu_0\sigma^2}{N\sigma_0^N+\sigma^2}\right)^2}{\frac{1}{\left(\frac{N}{\sigma^2}+\frac{1}{\sigma_0^2}\right)}} + \text{const.}$
3	465	9 行目	$\ln p(\boldsymbol{\Lambda} \mid \mathbf{X}) = \ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Lambda})$ $+ \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) - p(\boldsymbol{\mu} \mid \boldsymbol{\Lambda}, \mathbf{X}) - p(\mathbf{X}).$	$\ln p(\boldsymbol{\Lambda} \mid \mathbf{X}) = \ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Lambda})$ $+ \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) - \ln p(\boldsymbol{\mu} \mid \boldsymbol{\Lambda}, \mathbf{X}) - \ln p(\mathbf{X}).$
3	465	13-14 行目	$\ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Lambda})$ $= -\dfrac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}_n-\boldsymbol{\mu})$ $+\dfrac{N}{2}\lvert\boldsymbol{\Lambda}\rvert + \text{const.}$ $= -\dfrac{1}{2}\mathrm{Tr}\left(\sum_{n=1}^{N}(\mathbf{x}_n-\boldsymbol{\mu})(\mathbf{x}_n-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}\right)$ $+\dfrac{N}{2}\lvert\boldsymbol{\Lambda}\rvert + \text{const.}$	$\ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Lambda})$ $= -\dfrac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}(\mathbf{x}_n-\boldsymbol{\mu})$ $+\dfrac{N}{2}\ln\lvert\boldsymbol{\Lambda}\rvert + \text{const.}$ $= -\dfrac{1}{2}\mathrm{Tr}\left(\sum_{n=1}^{N}(\mathbf{x}_n-\boldsymbol{\mu})(\mathbf{x}_n-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Lambda}\right)$ $+\dfrac{N}{2}\ln\lvert\boldsymbol{\Lambda}\rvert + \text{const.}$

正誤表

3	465	下から 8-3 行目	$\begin{aligned} \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= -\frac{\beta_0}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Lambda}(\boldsymbol{\mu}-\boldsymbol{\mu}_0) \\ &\quad +\frac{1}{2}\lvert\boldsymbol{\Lambda}\rvert + \frac{\nu_0-D-1}{2}\lvert\boldsymbol{\Lambda}\rvert \\ &\quad -\frac{1}{2}\mathrm{Tr}(\mathbf{W}_0^{-1}\boldsymbol{\Lambda}) + \text{const.} \\ &= -\frac{1}{2}\beta_0\mathrm{Tr}\left((\boldsymbol{\mu}-\boldsymbol{\mu}_0)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Lambda}\right) \\ &\quad +\frac{1+\nu_0-D-1}{2}\lvert\boldsymbol{\Lambda}\rvert \\ &\quad -\frac{1}{2}\mathrm{Tr}(\mathbf{W}_0^{-1}\boldsymbol{\Lambda}) + \text{const.} \\ &= -\frac{1}{2}\mathrm{Tr}\left(\mathbf{W}_0^{-1}+\beta_0(\boldsymbol{\mu}-\boldsymbol{\mu}_0)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\right)\boldsymbol{\Lambda} \\ &\quad +\frac{\nu_0-D}{2}\lvert\boldsymbol{\Lambda}\rvert + \text{const.} \end{aligned}$	$\begin{aligned} \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) &= -\frac{\beta_0}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Lambda}(\boldsymbol{\mu}-\boldsymbol{\mu}_0) \\ &\quad +\frac{1}{2}\ln\lvert\boldsymbol{\Lambda}\rvert + \frac{\nu_0-D-1}{2}\ln\lvert\boldsymbol{\Lambda}\rvert \\ &\quad -\frac{1}{2}\mathrm{Tr}(\mathbf{W}_0^{-1}\boldsymbol{\Lambda}) + \text{const.} \\ &= -\frac{1}{2}\beta_0\mathrm{Tr}\left((\boldsymbol{\mu}-\boldsymbol{\mu}_0)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Lambda}\right) \\ &\quad +\frac{1+\nu_0-D-1}{2}\ln\lvert\boldsymbol{\Lambda}\rvert \\ &\quad -\frac{1}{2}\mathrm{Tr}(\mathbf{W}_0^{-1}\boldsymbol{\Lambda}) + \text{const.} \\ &= -\frac{1}{2}\mathrm{Tr}\left(\mathbf{W}_0^{-1}+\beta_0(\boldsymbol{\mu}-\boldsymbol{\mu}_0)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\right)\boldsymbol{\Lambda} \\ &\quad +\frac{\nu_0-D}{2}\ln\lvert\boldsymbol{\Lambda}\rvert + \text{const.} \end{aligned}$
3	469	6 行目 式	$-\frac{\beta}{2}\sum_{n=1}^{N}(\mathbf{w}^{\mathrm{T}}(\boldsymbol{\phi}(\mathbf{x}_n)-t_n)^2 + \text{const.}$	$-\frac{\beta}{2}\sum_{n=1}^{N}(\mathbf{w}^{\mathrm{T}}\boldsymbol{\phi}(\mathbf{x}_n)-t_n)^2 + \text{const.}$
3	469	8 行目	$p(\mathbf{w}, \beta) = \cdots$	$p(\mathbf{w}, \beta \mid \mathbf{t}) = \cdots$
3	472	16 行目	$\sigma'(x) = -\left(\frac{e^{-x}}{(1+e^{-x})^2}\right)\cdots$	$\sigma'(x) = -\left(\frac{1}{(1+e^{-x})^2}\right)\cdots$
3	485	8 行目 式	$L(\mathbf{w}, b, a_1, a_2) = \frac{1}{\lVert\mathbf{w}\rVert} - \cdots$	$L(\mathbf{w}, b, a_1, a_2) = \frac{1}{2}\lVert\mathbf{w}\rVert^2 - \cdots$
3	485	11 行目 式	$\mathbf{w} - a_1\mathbf{x}_1 - a_2\mathbf{x}_2 = 0,$	$\mathbf{w} - a_1\boldsymbol{\phi}(\mathbf{x}_1) - a_2\boldsymbol{\phi}(\mathbf{x}_2) = 0,$
3	485	14 行目 式	$\mathbf{w} = a_1(\mathbf{x}_1 - \mathbf{x}_2)$	$\mathbf{w} = a_1(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2))$
3	485	16 行目 式	$b = -\frac{1}{2}\mathbf{w}^{\mathrm{T}}(\mathbf{x}_1 + \mathbf{x}_2).$	$b = -\frac{1}{2}\mathbf{w}^{\mathrm{T}}(\boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)).$
3	485	18 行目 式	$b = -\frac{a_1}{2}(\mathbf{x}_1 - \mathbf{x}_2)^{\mathrm{T}}(\mathbf{x}_1 + \mathbf{x}_2).$	$b = -\frac{a_1}{2}(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2))^{\mathrm{T}}(\boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)).$
3	510	6 行目	$\ldots, 0 \le \theta < \pi.$	$\ldots, 0 \le \theta < 2\pi.$
3	511	1 行目 式	$\cdots = -\frac{1}{2}\left\{\left(\frac{1}{\sigma_X^2} + \frac{2}{\sigma_Y^2}\right)x^2 - \frac{2z}{\sigma_Y^2}x + \frac{z^2}{\sigma_Y^2}\right\}$	$\cdots = -\frac{1}{2}\left\{\left(\frac{1}{\sigma_X^2} + \frac{1}{\sigma_Y^2}\right)x^2 - \frac{2z}{\sigma_Y^2}x + \frac{z^2}{\sigma_Y^2}\right\}$

追記と注釈

1 巻・14 ページ・第 3 パラグラフ：「画像などの高次元データを入力とし，... 深層学習も例外ではない.」に対し，以下の脚注をつける.

「大規模な深層ネットワークでは，訓練データ数を一定としたもとで，パラメータを増やすと過学習がおき，パラメータ数とともに汎化誤差が増大するが，さらにパラメータ数を増やしていくと汎化誤差が減りはじめることがある（2 重下降とよばれる）．ただし，2 重下降がおこる原因は（いくかの仮説はあるものの）わかっておらず，この現象については研究途上である.」

1 巻・23 ページ・1.3.3 確率の解釈：頻度主義とベイズ主義 ：この節の冒頭に以下を追記.

「前述した公理で定義される確率を，現実の世界と結びつけ解釈する立場は，頻度主義とベイズ主義とに大きく 2 つにわかれる.」

1 巻・29 ページ・2 行目 ：追記.

具体的には，θ の確率密度関数は，$p(\theta) \propto \theta^3$ と $\int_0^1 p(\theta)\, d\theta = 1$ から $p(\theta) = 4\theta^3, 0 \le \theta \le 1,$ である．また，先の例にあげた線形回帰モデルのパラメータ w の平均 0 のガウス事前分布は，w の事後分布の平均が 0 から極端にはなれないようにするいかりの役割をはたす.

1 巻・32 ページ・下から 8 行目から 7 行目 ：「正規化係数」に以下のカッコ書きを追記.

「正規化係数（正規化定数の逆数）」

1 巻・34 ページ・下から 4 行目 (脚注を除く） ：「さらに，データ $(\mathbf{x}, t)$ がしたがう」に追記し以下とする.

「さらに，ここでは，$\mathbf{x}$ も確率変数（さきの線形回帰モデルでは通常の変数とした）として，データ $(\mathbf{x}, t)$ がしたがう」

1 巻・92 ページ・図 3.4 キャプション「を示す.」の後に以下を追記.

(d) の左端の図では，2 本の直線が，表示範囲をはずれているため表示されていない．また，表示されている 5 本の直線がすべて右肩下がりなのは偶然である.

1 巻・121 ページ 9 行目：「初期値」の後に「，たとえば,」を追記.

「すなわち，初期値，たとえば，$\mathbf{w}_0 = \mathbf{0}$ から出発し，$\cdots$」

1 巻・147-149 ページ・付記 ：凸関数に関する定理の証明.

この項は，純粋に数学的事項なので，文字フォントのポイントを 1 段小さくすべきところでした.

1 巻・199 ページ・2 行目：「である.」の後に以下を追記.

「同様に，$\dfrac{d\mathbf{a}^{\mathrm{T}}}{d\mathbf{b}} \equiv \left(\dfrac{d\mathbf{a}}{d\mathbf{b}^{\mathrm{T}}}\right)^{\mathrm{T}}$ である.」

1 巻・199 ページ・4 行目：式 C1) の後に以下を追記.

「C1’) $\dfrac{d}{d\mathbf{x}}(\mathbf{A}\mathbf{x})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}}.$」

1 巻・199 ページ・5 行目：式 C2) の後に以下を追記.

「C2’) $\dfrac{d\mathbf{x}^{\mathrm{T}}}{d\mathbf{x}} = \mathbf{I}^{\mathrm{T}} = \mathbf{I}.$」

1 巻・200 ページ・1 行目：「ヘッセ行列といわれる.」の後に以下を追記.

「本書では，簡単のため，$\nabla(\nabla f(\mathbf{x}))^{\mathrm{T}}$ を $\nabla\nabla f(\mathbf{x})$ と略記する.」

追記と注釈

2 巻・292 ページ・下から 9 行目：「$\cdots$ ならべた行列である.」の後に以下を追記.
「また，$\mathbf{z}_n = \mathbf{W}^{\mathrm{T}}\mathbf{x}_n$ である.」
3 巻・445 ページ・7 行目：「である.」の後に以下を追記.
「同様に，$\dfrac{d\mathbf{a}^{\mathrm{T}}}{d\mathbf{b}} \equiv \left(\dfrac{d\mathbf{a}}{d\mathbf{b}^{\mathrm{T}}}\right)^{\mathrm{T}}$ である.」
3 巻・445 ページ・9 行目：式 C1) の後に以下を追記.
「C1') $\dfrac{d}{d\mathbf{x}}(\mathbf{A}\mathbf{x})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}}$.」
3 巻・445 ページ・10 行目：式 C2) の後に以下を追記.
「C2') $\dfrac{d\mathbf{x}^{\mathrm{T}}}{d\mathbf{x}} = \mathbf{I}^{\mathrm{T}} = \mathbf{I}$.」
3 巻・446 ページ・2 行目：「ヘッセ行列といわれる.」の後に以下を追記.
「本書では，簡単のため，$\nabla(\nabla f(\mathbf{x}))^{\mathrm{T}}$ を $\nabla\nabla f(\mathbf{x})$ と略記する.」