

正誤情報

このたびは森北出版株式会社発行の書籍をお買い求めいただき、誠にありがとうございました。下記の書籍につきまして誤りのある箇所がございましたので、お詫びし訂正させていただきます。

2020年11月17日 森北出版株式会社 生産マネジメント部

タイトル

はじめてのパターン認識

正誤対象

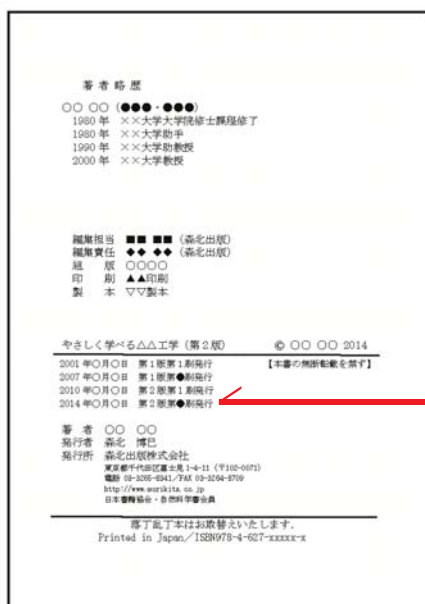
お手持ちの書籍の刷数をお調べのうえ、下の表をご覧ください。正誤表内の一番左に「対応刷数」という列がございます。該当する刷数の訂正情報をご参照下さい。

なお、刷数につきましては下記「刷数の調べ方」をご参照ください。

お持ちの本の刷数	
1	対応刷数 1 より 5 までをご参照ください(別ファイル6刷以降用もご参照ください)
2	対応刷数 2 より 5 までをご参照ください(別ファイル6刷以降用もご参照ください)
3	対応刷数 3 より 5 までをご参照ください(別ファイル6刷以降用もご参照ください)
4	対応刷数 4 より 5 までをご参照ください(別ファイル6刷以降用もご参照ください)
5	対応刷数 5 をご参照ください(別ファイル6刷以降用もご参照ください)
それ以降	6刷以降用の正誤情報が別ファイルにありますのでそちらもご参照ください

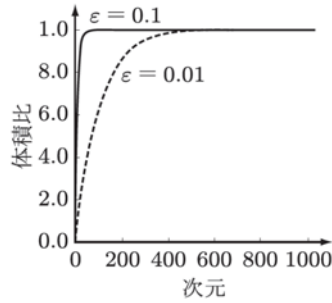
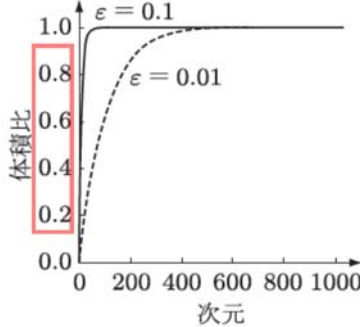
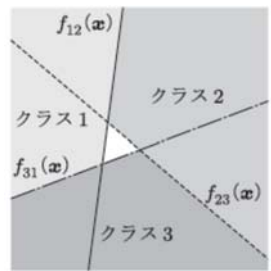
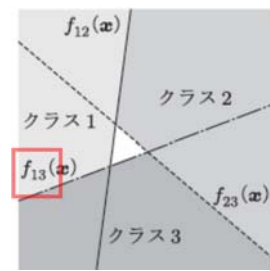
刷数の調べ方

本の一番後ろのページ(広告等除く)に下図のようなページがございます。ご参照いただき、お持ちの本の刷数をお調べください。



日付の最も新しい行に記載された数字がお持ちの本の刷数となります

対応刷数	頁	行数, 図・表・式番号	誤	正
3	9	図 2.1(a)	p (4 か所)	P
5	10	15 行目	…場合は, ダミー変数表現を…	…場合は, <u>クラス数のビット幅をもった</u> ダミー変数表現を…
4	16	下から	…わかりやすい回帰の例を…	…わかりやすい回帰 (関数近似) の例を…
5		2 行目	…わかりやすい関数近似の例を…	
5	22	11 行目	…尤度の和は $\sum_{i=1}^K p(\mathbf{x} C_i) \neq 1$ なので, 厳密にいうと…	…尤度の和は $\sum_{i=1}^K p(\mathbf{x} C_i) \neq 1$ なので, <u>逆に見た尤度は</u> 厳密にいうと…
2	28	10 行目	で表される.識別規則は,	で表される.識別規則は <u>式(3.13)にしたがって</u> ,
5	30	下から 3 行目	confusion marix	confusion matrix
5	31	8 行目	偽陽性率は偽のものを真と判断した割合, 真陽性率は真のものを正しく真と判断した割合である.	偽陽性率は陰性のものを陽性と判断した割合, 真陽性率は陽性のものを正しく陽性と判断した割合である.
4	32	下から 8 行目	…濃い網かけの部分 ε_2 が陽性…	…陰性のクラスのうち, 濃い網かけの部分 ε_2 が陽性…
5	32	下から 7 行目	を第 1 種の誤り (miss), ε_2 を第 2 種の誤り (false alarm) …	を第 2 種の誤り (miss), ε_2 を第 1 種の誤り (false alarm) …
2	33	5 行目	ROC 曲線下面積(AUR:area under ROC curve)	ROC 曲線下面積(AUC:area under an ROC curve)
2	33	7,8,9 行目	AUR (3 箇所)	AUC
1	35	章末問題 3.1(2)	誤りの期待値を求めよ.	<u>ベイズ誤り率</u> を求めよ.
2	37	式(4.2)	右辺 $= \int (x_i - \mu_i)(x_j - \mu_j) p(x_i, x_j) dx_i dx_j$	$= \iint (x_i - \mu_i)(x_j - \mu_j) p(x_i, x_j) dx_i dx_j$
2	40	下から 2 行目	行列 \mathbf{S} は直交行列となるので, …	行列 \mathbf{S} は <u>正規</u> 直交行列となるので, …
2	41	1 行目	直交行列 \mathbf{S} では, …	<u>正規</u> 直交行列 \mathbf{S} では, …
2	42	図 4.3		大文字の \mathbf{S} を小文字の \mathbf{s} にする(6 箇所)

5	47	7~8行目	で定義された2次元正規分布から100個のデータをサンプルし、確率の等高線と共にグラフに示す.	で定義された2次元正規分布からサンプルされた100個のデータを、確率の等高線と共に示した.
2	52	式(4.40) 1行目 右辺	$\dots = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$	$\dots = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$
2	53	式(4.43)	$\frac{\partial \mathcal{L}(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{1}{N} \frac{1}{\sigma^2} \dots$	$\frac{\partial \mathcal{L}(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{N}{2} \frac{1}{\sigma^2} + \dots$
5			$\frac{\partial \mathcal{L}(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{N}{2} \frac{1}{\sigma^2} \dots$	
3	54	式(5.1)	識別クラス = $\begin{cases} \arg \min_i d(\mathbf{x}, \mathbf{x}_j^{(i)}) \\ \text{リジェクト} \end{cases}$	識別クラス = $\left\{ \arg \min_i \left\{ \min_j d(\mathbf{x}, \mathbf{x}_j^{(i)}) \right\} \right\}$ リジェクト
2	57	8行目	…6500個のテストデータのうち、…	…6500個の <u>全</u> テストデータのうち、…
3	64	式(5.11)	$S_d = \frac{2\pi^{1/2}}{\Gamma(d/2)}$	$S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$
2	64	図 5.11 (b)		
2	69	下から 10行目	…BBD(balanced	…BBD(balanced
2	73	図 6.2(b)		

5	74	12行目	を満たす $K-1$ 個の識別境界が...	を満たす K 個の識別境界が...
2	76	式(6.11)	$E(w) = \sum_{i=1}^N (t_i - f(x_i))^2 = (t - Xw)^T (t - Xw)$	$E(w) = \sum_{i=1}^N (t_i - f(x_i))^2 = (t - Xw)^T (t - Xw) = t^T t - 2t^T Xw + w^T X^T Xw$
2	76	式(6.12)	$\frac{\partial E(w)}{\partial w} = -2X^T(t - Xw) = 0$	$\frac{\partial E(w)}{\partial w} = -2X^T t + 2X^T Xw = 0$
1	76	脚注	…スカラー関数のベクトルによる微分は、付録 A.2 に解説しているので参照されたい。	…スカラー関数のベクトルによる微分は、付録 A.1 に解説しているので参照されたい。
4	77	最下行	同時座標表現では...	同次座標表現では...
2	78	式(6.16)	$G(x) = \hat{W}^T x = (w_1, \dots, w_K)^T x = (g_1(x), \dots, g_K(x))^T$	$f(x) = \hat{W}^T x = (w_1, \dots, w_K)^T x = (f_1(x), \dots, f_K(x))^T$
2	78	式(6.17)	識別クラス = $\arg \max_j g_j(x)$	識別クラス = $\arg \max_j f_j(x)$
2	80	図 6.6	<p>Figure 6.6 illustrates the relationship between margin size and error rate. On the left, a linear decision boundary separates two classes, C_1 and C_2. The margins are labeled S_1 and S_2, and the decision boundary is w_0. The error rate is low (誤りが少ない). On the right, the margins are larger, labeled S'_1 and S'_2, and the decision boundary is w'_0. The error rate is high (誤りが多い).</p>	<p>Figure 6.6 illustrates the relationship between margin size and error rate. On the left, a linear decision boundary separates two classes, C_1 and C_2. The margins are labeled S_1 and S_2, and the decision boundary is w_0. The error rate is low (誤りが少ない). On the right, the margins are larger, labeled S'_1 and S'_2, and the decision boundary is w'_0. The error rate is high (誤りが多い).</p>
2	82	式(6.33)	$2 \left(\frac{\partial h}{\partial \sigma_1^2} \sum_1 + \frac{\partial h}{\partial \sigma_2^2} \sum_2 \right) w = \frac{\partial h}{\partial m_1} \mu_1 + \frac{\partial h}{\partial m_2} \mu_2$	$2 \left(\frac{\partial h}{\partial \sigma_1^2} \sum_1 + \frac{\partial h}{\partial \sigma_2^2} \sum_2 \right) w = \left(\frac{\partial h}{\partial m_1} \mu_1 + \frac{\partial h}{\partial m_2} \mu_2 \right)$
2	83	式(6.39)	$\dots = \frac{P(C_k)(P(C_1)(m_1 - \bar{m})^2 + P(C_2)(m_2 - \bar{m})^2)}{(P(C_1)\sigma_1^2 + P(C_2)\sigma_2^2)^2}$	$\dots = \frac{P(C_k)(P(C_1)(m_1 - \bar{m})^2 + P(C_2)(m_2 - \bar{m})^2)}{(P(C_1)\sigma_1^2 + P(C_2)\sigma_2^2)^2}$
2	83	式(6.42)	$w_0 = -w^T (P(C_1)\mu_1 + \dots)$	$w_0 = \bar{m} - w^T (P(C_1)\mu_1 + \dots)$

4	86	12 行目	$\mathbf{y} = (y_1, \dots, y_{K-1}), \dots$	$\mathbf{y} = (y_1, \dots, y_{K-1})^T, \dots$
5	86	脚注	行列のランクと写像の次元の関係については第 9 章でも触れる。	行列のランクと写像の次元の関係については第 9 章でも触れる。 <u>行列</u> <u>$(\boldsymbol{\mu}_k - \boldsymbol{\mu})(\boldsymbol{\mu}_k - \boldsymbol{\mu})^T$ のランクが 1 になることが、例題 9.3 に示されてい</u> <u>る。</u>
4	87	式 (6.56)	$\dots = \text{Tr}(\mathbf{S}_W^{-1} \mathbf{S}_B) = \dots$	$\dots = \text{Tr}(\tilde{\mathbf{S}}_W^{-1} \tilde{\mathbf{S}}_B) = \dots$
2	89	下から 2 行目 ～ 式(6.61)	N 人の喫煙量 $\hat{\mathbf{x}} = (x_1, \dots, x_N)^T$ を観測したとき、肺がんになる確率 を、 $p(1 x_1, \dots, x_N) = f(\mathbf{x}) = \frac{1}{1 + \exp(-(w_0 + w_1 x_1 + \dots + w_N x_N))}$	N 人の喫煙量 $x_i, i=1, \dots, N$ を観測したとき、 <u>喫煙量 x の人が</u> 肺がんになる 確率を、 $P(1 x) = f(x) = \frac{1}{1 + \exp(-(w_0 + w_1 x))}$
2	90	1～2 行目	で表す。 $\mathbf{w} = (w_0, w_1, \dots, w_N)^T$ とし、 $\hat{\mathbf{x}}$ にはバイアス項に対応す る 1 を追加して $\mathbf{x} = (1, \hat{\mathbf{x}}^T)^T$ とする。	で表す。 $\mathbf{w} = (w_0, w_1)^T$ とし、 <u>x</u> にはバイアス項に対応する 1 を追加して $\mathbf{x} = (1, x)^T$ とする。
2	90	式(6.63)	$a = \ln \frac{p(1 \mathbf{x})}{1 - p(1 \mathbf{x})} = \mathbf{w}^T \mathbf{x}$	$a = \ln \frac{P(1 x)}{1 - P(1 x)} = \mathbf{w}^T \mathbf{x}$
2	90	式(6.64)	$\frac{p(1 \mathbf{x})}{1 - p(1 \mathbf{x})} = \frac{p(1 \mathbf{x})}{p(0 \mathbf{x})} = \exp(\mathbf{w}^T \mathbf{x})$	$\frac{P(1 x)}{1 - P(1 x)} = \frac{P(1 x)}{P(0 x)} = \exp(\mathbf{w}^T \mathbf{x})$
2	90	14 行目	\mathbf{x} の中の x_1 が 1 増えた状態 $\tilde{\mathbf{x}} = (1, (x_1 + 1), x_2, \dots, x_N)$ を考えよう。	\mathbf{x} の中の <u>x</u> が 1 増えた状態 <u>$\tilde{\mathbf{x}} = (1, (x_1 + 1))$</u> を考えよう。 (再修正あり。対応刷数 4 もご参照下さい)
4	90	14 行目	\mathbf{x} の中の x が 1 増えた状態 $\tilde{\mathbf{x}} = (1, (x_1 + 1))$ を考えよう。	\mathbf{x} の中の x が 1 増えた状態 $\tilde{\mathbf{x}} = (1, (x_1 + 1))^T$ を考えよう。

2	90	式(6.65)	右のように修正	$\frac{P(1 \tilde{\mathbf{x}})}{1 - P(1 \tilde{\mathbf{x}})} = \frac{\exp(w_0 + w_1(x_1 + 1))}{\exp(w_0 + w_1 x)} = \exp(w_1)$
				(再修正あり。対応刷数4もご参照下さい)
4	90	式(6.65)	$\dots = \frac{\exp(w_0 + w_1(x_1 + 1))}{\exp(w_0 + w_1 x)} = \dots$	$\dots = \frac{\exp(w_0 + w_1(x + 1))}{\exp(w_0 + w_1 x)} = \dots$
2	90	下から 11行目	となる。すなわち、 x_1 が1単位分増えることによる…	となる。すなわち、 x が1単位分増えることによる…
2	91	表 6.1 オッズ比	$P(B)/(1-P(B))/P(A)/(1-P(A))$	$(P(B)/(1-P(B)))/(P(A)/(1-P(A)))$
2	91,92	6.4.3節 全体	α (21箇所)	π
2	91	3~4行目	… t が1となる確率を $p(t=1)=\alpha$ で、0となる確率を $p(t=0)=1-\alpha$ で表す…	… t が1となる確率を $P(t=1)=\pi$ で、0となる確率を $P(t=0)=1-\pi$ で表す。 …
2	91	下から 3行目	が得られる。	が得られる。 $\underline{x_i}$ が入力ベクトル、 $\underline{t_i}$ がそのクラスを表す教師信号である。
2	91	最下行	…対数尤度関数を…	… <u>負</u> の対数尤度関数を…
2	92	式(6.74)	$\phi(\mathbf{x}) = (\phi_0 = 1, \phi_1(\mathbf{x}), \dots, \phi_M(\mathbf{x}))^T$	$\varphi(\mathbf{x}) = (\phi_0 = 1, \phi_1(\mathbf{x}), \dots, \phi_M(\mathbf{x}))^T$ (φ はベクトル記号)
2	92	式(6.75)	$a_k = \mathbf{w}_k^T \phi(\mathbf{x})$	$a_k = \mathbf{w}_k^T \varphi(\mathbf{x})$ (φ はベクトル記号)
2	92	下から 3行目	…ほかはすべて0のダミー符号化表現とする。	…ほかはすべて0のダミー <u>変数</u> 表現とする。
2	92	式(6.76)	$p(T \mathbf{w}_1, \dots, \mathbf{w}_K) = \dots$	$P(T \mathbf{w}_1, \dots, \mathbf{w}_K) = \dots$
2	93	式(6.78)	$E(\mathbf{w}_1, \dots, \mathbf{w}_K) = -\ln p(T \mathbf{w}_1, \dots, \mathbf{w}_K) = -\sum_{i=1}^N \sum_{k=1}^K t_{ik} \pi_{ik}$	$E(\mathbf{w}_1, \dots, \mathbf{w}_K) = -\ln P(T \mathbf{w}_1, \dots, \mathbf{w}_K) = -\sum_{i=1}^N \sum_{k=1}^K t_{ik} \ln \pi_{ik}$
2	93	8行目	$\frac{\partial E}{\partial \mathbf{w}_j} = \dots = -\sum_{i=1}^N \sum_{k=1}^K t_{ik} \frac{1}{\pi_{ik}} (\delta_{ij} - \pi_{ij}) \mathbf{x}_i$	$\frac{\partial E}{\partial \mathbf{w}_j} = \dots = -\sum_{i=1}^N \sum_{k=1}^K t_{ik} (\delta_{kj} - \pi_{ij}) \mathbf{x}_i$

2	93	10~11行目	<p>が得られる. ここで, δ_{ij} はクロネッカーのデルタで,</p> $\delta_{ij} = \begin{cases} 1 & (i=j \text{ の場合}) \\ 0 & (i \neq j \text{ の場合}) \end{cases}$	<p>が得られる. ここで, δ_{kj} はクロネッカーのデルタで,</p> $\delta_{kj} = \begin{cases} 1 & (k=j \text{ の場合}) \\ 0 & (k \neq j \text{ の場合}) \end{cases}$																																								
2	98	最下行	$\mathbf{w}^T \mathbf{x} = 0$ に対して取れるマージンの大きさ D は, ...	$\mathbf{w}^T \mathbf{x}$ に対して取れるマージンの大きさ $D(\mathbf{w})$ は, ...																																								
2	99	式(7.2)	$\rho = \dots$	$\rho(\mathbf{w}) = \dots$																																								
2	99	3行目	... ρ をクラス間マージンとよぶ.	... $\rho(\mathbf{w})$ をクラス間マージンとよぶ.																																								
5	99	脚注	...超平面向きの決まらない. $d+1$ 個のベクトルが...	...超平面向きの決まらない. <u>同次座標系</u> で $d+1$ 個のベクトルが...																																								
2	100	11行目	ろを通る. 識別関数からそれぞれの...	ろを通る. 識別 <u>境界</u> からそれぞれの...																																								
5	100	下から9行目	$\mathbf{w} = b(-2/3, -2/3)^T$	$\mathbf{w} = b(-2/3, -2/3)^T$																																								
2	102	表 7.2	<table border="1"> <thead> <tr> <th>x_1</th> <th>x_2</th> <th>出力</th> <th>教師データ</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> <td>0</td> <td>-1</td> </tr> <tr> <td>0</td> <td>1</td> <td>1</td> <td>+1</td> </tr> <tr> <td>0</td> <td>1</td> <td>1</td> <td>+1</td> </tr> <tr> <td>1</td> <td>1</td> <td>0</td> <td>-1</td> </tr> </tbody> </table>	x_1	x_2	出力	教師データ	0	0	0	-1	0	1	1	+1	0	1	1	+1	1	1	0	-1	<table border="1"> <thead> <tr> <th>x_1</th> <th>x_2</th> <th>出力</th> <th>教師データ</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> <td>0</td> <td>-1</td> </tr> <tr> <td>0</td> <td>1</td> <td>1</td> <td>+1</td> </tr> <tr> <td>1</td> <td>0</td> <td>1</td> <td>+1</td> </tr> <tr> <td>1</td> <td>1</td> <td>0</td> <td>-1</td> </tr> </tbody> </table>	x_1	x_2	出力	教師データ	0	0	0	-1	0	1	1	+1	1	0	1	+1	1	1	0	-1
x_1	x_2	出力	教師データ																																									
0	0	0	-1																																									
0	1	1	+1																																									
0	1	1	+1																																									
1	1	0	-1																																									
x_1	x_2	出力	教師データ																																									
0	0	0	-1																																									
0	1	1	+1																																									
1	0	1	+1																																									
1	1	0	-1																																									
4	107	式 (7.19)	$\Delta w_{jk}(\tau) = \eta \sum_{n=1}^N \delta_j^n x_k^n$	$\Delta w_{ji}(\tau) = \eta \sum_{n=1}^N \delta_j^n x_i^n$																																								
5	110	式(7.22)	$\Delta w_{kj}(\tau) = \sum_{n=1}^N \left(-\eta \frac{\partial E_n(\mathbf{w}) + \lambda R(\mathbf{w})}{\partial w_{kj}} \right) = \eta \sum_{n=1}^N \delta_k^n V_j^n - 2\lambda w_{kj}$	$\Delta w_{kj}(\tau) = -\eta \frac{\partial \tilde{E}(\mathbf{w})}{\partial w_{kj}} = -\eta \left(\sum_{n=1}^N \frac{\partial E_n(\mathbf{w})}{\partial w_{kj}} + \lambda \frac{\partial R(\mathbf{w})}{\partial w_{kj}} \right)$ $= \eta \sum_{n=1}^N \delta_k^n V_j^n - 2\eta \lambda w_{kj}$																																								
2	112	式(7.23)	$E(\mathbf{w}) = -\sum_{n=1}^N \sum_{k=1}^K t_k^n \log \sigma_k^n$	$E(\mathbf{w}) = -\sum_{n=1}^N \sum_{k=1}^K t_k^n \ln \sigma_k^n$																																								

5	115	式(8.4)	$\dots = \min_{\mathbf{x} \in C_{y=+1}} \frac{\mathbf{w}^T \mathbf{x}}{\ \mathbf{w}\ } - \max_{\mathbf{x} \in C_{y=-1}} \frac{\mathbf{w}^T \mathbf{x}}{\ \mathbf{w}\ } = \dots$	$\dots = \min_{\mathbf{x}_i \in C_{t_i=+1}} \frac{\mathbf{w}^T \mathbf{x}_i}{\ \mathbf{w}\ } - \max_{\mathbf{x}_i \in C_{t_i=-1}} \frac{\mathbf{w}^T \mathbf{x}_i}{\ \mathbf{w}\ } = \dots$
2	115	図 8.1		
2	119	図 8.3		
2	120	式(8.17)	$\tilde{L}(\mathbf{w}, b, \alpha, \xi, \mu) = \dots$	$\tilde{L}_p(\mathbf{w}, b, \alpha, \xi, \mu) = \dots$
2	120	下から 7 行目	(1) $\frac{\partial \tilde{L}}{\partial \mathbf{w}} \Big _{\mathbf{w}=\mathbf{w}_0} = \dots$	(1) $\frac{\partial \tilde{L}_p}{\partial \mathbf{w}} \Big _{\mathbf{w}=\mathbf{w}_0} = \dots$
2	120	下から 6 行目	(2) $\frac{\partial \tilde{L}}{\partial b} = \dots$	(2) $\frac{\partial \tilde{L}_p}{\partial b} = \dots$
2	120	下から 5 行目	(3) $\frac{\partial \tilde{L}}{\partial \xi_i} = \dots$	(3) $\frac{\partial \tilde{L}_p}{\partial \xi_i} = \dots$
2	121	双対問題[2]	制約条件: $0 \leq \alpha_i \leq C, \quad \mathbf{a}^T \mathbf{t} = 0$	制約条件: $0 \leq \alpha_i \leq C, \quad \mathbf{a}^T \mathbf{t} = 0$
2	122	双対問題[3]		

2	123	式(8.25)	$\phi(\mathbf{u}) = \dots$	$\varphi(\mathbf{u}) = \dots$ (φ はベクトル記号)
2	123	式(8.26)	$\phi(\mathbf{v}) = \dots$	$\varphi(\mathbf{v}) = \dots$ (φ はベクトル記号)
2	123	下から8行目	のように分離すれば, $K_2(\mathbf{u}, \mathbf{v}) = \phi(\mathbf{u})^T \phi(\mathbf{v})$ と表現できる...	のように分離すれば, $K_2(\mathbf{u}, \mathbf{v}) = \varphi(\mathbf{u})^T \varphi(\mathbf{v})$ と表現できる... (φ はベクトル記号)
5	125	最下行	...広がりを表すパラメータを $\sigma = 0.2, 0.4, 0.8, \underline{2.0}, 4.0$...広がりを表すパラメータを $\sigma = 0.2, 0.4, 0.8, 4.0$
2	126	図 8.4 キャプション	...200 個のデータの一個抜き法による訓練誤差(実線)と汎化誤差(点線)を,200 個のデータの訓練誤差(実線)と一個抜き法による汎化誤差(点線)を, ...
3	126	図 8.4		<p>Figure 8.4: A line graph showing error vs. C for different sigma values. The x-axis is C (1, 10, 100, 1000, 10000) on a log scale. The y-axis is error (0.0 to 0.5). Training error (solid lines) decreases as C increases. Test error (dotted lines) initially increases and then levels off. Higher sigma values result in higher test errors.</p>
2	129	10 行目	(1) $\left. \frac{\partial \tilde{L}}{\partial \mathbf{w}} \right _{\mathbf{w}=\mathbf{w}_0} = \dots$	(1) $\left. \frac{\partial \tilde{L}_p}{\partial \mathbf{w}} \right _{\mathbf{w}=\mathbf{w}_0} = \dots$
2	129	11 行目	(2) $\frac{\partial \tilde{L}}{\partial b} = \dots$	(2) $\frac{\partial \tilde{L}_p}{\partial b} = \dots$
2	129	12 行目	(3) $\frac{\partial \tilde{L}}{\partial \xi_i} = \dots$	(3) $\frac{\partial \tilde{L}_p}{\partial \xi_i} = \dots$

2	129	13 行目	(4) $\frac{\partial \tilde{L}}{\partial \rho} = \dots$	(4) $\frac{\partial \tilde{L}_p}{\partial \rho} = \dots$
2	137	下から 6 行目	N 個のデータ \mathbf{x}_i を...	N 個のデータ $\mathbf{x}_i - \bar{\mathbf{x}}$ を...
2	137	下から 3~2 行目	<p>が得られる. この変換後のデータの分散は,</p> $\text{Var}\{\mathbf{s}_j\} \propto \mathbf{s}_j^T \mathbf{s}_j = (\bar{\mathbf{X}} \mathbf{a}_j)^T \bar{\mathbf{X}} \mathbf{a}_j = \mathbf{a}_j^T \bar{\mathbf{X}}^T \bar{\mathbf{X}} \mathbf{a}_j = \mathbf{a}_j^T \text{Var}\{\bar{\mathbf{X}}\} \mathbf{a}_j$	<p>が得られる. この変換後のデータの分散は, $\mathbf{s}_{1j}, \dots, \mathbf{s}_{Nj}$ の平均が 0 となる ので,</p> $\text{Var}\{\mathbf{s}_j\} = \frac{1}{N} \mathbf{s}_j^T \mathbf{s}_j = \frac{1}{N} (\bar{\mathbf{X}} \mathbf{a}_j)^T (\bar{\mathbf{X}} \mathbf{a}_j) = \frac{1}{N} \mathbf{a}_j^T \bar{\mathbf{X}}^T \bar{\mathbf{X}} \mathbf{a}_j = \mathbf{a}_j^T \text{Var}\{\bar{\mathbf{X}}\} \mathbf{a}_j$
5	139	11 行目	...データ行列のランクが 10 から 9 になる.	...データ行列の最大ランクが 10 から 9 になる.
3	141	5 行目	高い精度再構成できていることが	高い精度で再構成できている
5	144 145	例題 9.3 解答	右のように修正	$\mathbf{u} = (u_1, \dots, u_d)^T$ とする. $\mathbf{U} = \mathbf{u}\mathbf{u}^T = (u_1\mathbf{u}, \dots, u_d\mathbf{u})$ なので, 行列 \mathbf{U} の列ベクトル空間を張るのはベクトル \mathbf{u} のみとなり, 行列 \mathbf{U} のランクは 1 になる. \mathbf{U} の i, j 成分は $U_{ij} = u_i u_j$, j, i 成分は $U_{ji} = u_j u_i$ となるので, 対称行列になる. 射影行列は対称行列の和なので, 対称行列になる.
2	146	1 行目	$= \sum_{j=1}^{d_i} \mathbf{u}_{ij}^T E\{\mathbf{x}^T \mathbf{x} \mathbf{x} \in C_i\} \mathbf{u}_{ij} -$	$= \sum_{j=1}^{d_i} \mathbf{u}_{ij}^T E\{\mathbf{x}\mathbf{x}^T \mathbf{x} \in C_i\} \mathbf{u}_{ij} -$
5	147	下から 7 行目	...そろえた 650 個の学習データ...	...そろえた各クラス 650 個の学習データ...
5	165	10.4.1 の 2 行目	$\dots = \frac{1}{(2\pi)^{d/2} \Sigma_k ^{1/2}} \dots$	$\dots = \frac{1}{(2\pi)^{d/2} \Sigma_k ^{1/2}} \dots$
5	168	式(10.22)	$= \frac{p(z_{ik} = 1)p(\mathbf{x}_i z_{ik} = 1)}{\sum_{j=1}^K p(z_{ij} = 1)p(\mathbf{x}_i z_{ij} = 1)} = \frac{\pi_k N(\mathbf{x}_i \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j N(\mathbf{x}_i \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} = \gamma(z_{ik})$	$= \frac{p(z_{ik} = 1)p(\mathbf{x}_i z_{ik} = 1)}{\sum_{j=1}^K p(z_{ij} = 1)p(\mathbf{x}_i z_{ij} = 1)} = \frac{\pi_k N(\mathbf{x}_i \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j N(\mathbf{x}_i \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} = \gamma(z_{ik})$
2	170	8 行目	$= \sum_{i=1}^N \gamma(z_{ik}) \sum_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) = 0$	$= \sum_{i=1}^N \gamma(z_{ik}) \sum_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) = 0$

5	173	3行目	…，更新により Q 関数の値はつねに増加するか…	…，更新ごとに Q 関数を最大化すれば，尤度は必ず増加するか…
2	173	9行目	vesicolor の 1 個と…	versicolor の 1 個と…
2	182	式(11.9)	$\dots = -\sum_{i=1}^K P(C_i t) \log P(C_i t)$	$\dots = -\sum_{i=1}^K P(C_i t) \ln P(C_i t)$
3	182	式(11.10)	右のように修正	$\begin{aligned} \mathcal{I}(t) &= \sum_{i=1}^K \sum_{j \neq i} P(C_i t) P(C_j t) = \left(\sum_{i=1}^K P(C_i t) \right)^2 - \sum_{i=1}^K P^2(C_i t) \\ &= 1 - \sum_{i=1}^K P^2(C_i t) = \sum_{i=1}^K P(C_i t) (1 - P(C_i t)) \end{aligned}$
3	182	3行目	$\sum_{j \neq i} P(C_i t) P(C_j t)$	$\sum_{i=1}^K \sum_{j \neq i} P(C_i t) P(C_j t)$
2	182	下から 4行目	$\mathcal{I}(1) = 100/150 \cdot 50/150 = 2/9$ である.	$\mathcal{I}(1) = 100/150 \cdot 50/150 + 50/150 \cdot 100/150 = (100/150 \cdot 50/150) \times 2 = 4/9$ である.
2	182	下から 2行目	$\angle \mathcal{I}(\text{PL} < 2.45, 1) = \frac{2}{9} - \left(\frac{100}{150} \frac{50}{100} \frac{50}{100} + \frac{50}{100} \frac{0}{50} \frac{50}{50} \right) = 0.0556$	$\angle \mathcal{I}(\text{PL} < 2.45, 1) = \frac{4}{9} - \left(\frac{100}{150} \left(\frac{50}{100} \frac{50}{100} \right) \times 2 + \frac{50}{150} \left(\frac{0}{50} \frac{50}{50} \right) \times 2 \right) = \frac{1}{9} = 0.111$
2	183	例題 11.1 解答 1行目	$\angle \mathcal{I}(\text{PL} < 3.1, 1) = \frac{2}{9} - \left(\frac{99}{150} \frac{49}{99} \frac{50}{99} + \frac{51}{150} \frac{1}{51} \frac{50}{51} \right) = 0.0507$	$\angle \mathcal{I}(\text{PL} < 3.1, 1) = \frac{4}{9} - \left(\frac{99}{150} \left(\frac{49}{99} \frac{50}{99} \right) \times 2 + \frac{51}{150} \left(\frac{1}{51} \frac{50}{51} \right) \times 2 \right) = 0.101$
4	184	下から 6行目	もし， $R_\alpha(T_t) < R(t)$ であれば，	もし， $R_\alpha(T_t) < R_\alpha(t)$ であれば，
5	185	アルゴリズム 11.1	(1) $T^0 = T, k = 0$ とする.	(1) $T^0 = T, \alpha_0 = 0, k = 0$ とする.
5	185	アルゴリズム 11.1	(2)の後に(3)として右の行を挿入し，(3)以下はそれぞれ 1 つずつ番号をずらす	(3) $\alpha_{k+1} = g_k(t^k)$ とする.
5	185	下から 4行目	$g(7) = \alpha_1 = \dots$	$\alpha_1 = g(7) = \dots$
5	186	1行目	…ノード 4 の子孫を制定した結果を…	…ノード 4 の子孫を <u>剪定</u> した結果を…
2	187	図 11.5	木のサイズによる再代入誤りと交差検証法による誤り	木のサイズによる再代入誤り率と， <u>交差確認法による誤り率の変化</u>

2	187	例題 11.2	図 11.5 の $g(4)$ の値を求めよ。	図 11.4(b) の $g(4)$ の値を求めよ。
3	190	4 行目	$\alpha_m > 1$ となり,	$\alpha_m > 0$ となり,
2	192	式(11.32)	$w_i^{m+1} = w_i^m \exp\left(-\frac{\alpha_m}{2}\right) \exp(I(y_m(\mathbf{x}_i) \neq t_i))$	$w_i^{m+1} = w_i^m \exp\left(-\frac{\alpha_m}{2}\right) \exp(\alpha_m I(y_m(\mathbf{x}_i) \neq t_i))$
2	192	下から 11 行目	…式(11.2)を最小にする…	…式(11.21)を最小にする…
2	195	最下行	の値 $f(\mathbf{x}_i^{(x)})$ を, …	の値 $f_k(\mathbf{x})$ を, …
2	196	式(11.36)	$f(\mathbf{x}_i^{(x)}) = \sum_{i=1}^N \left(\log p_k(\mathbf{x}_i^{(x)}) - \frac{1}{K} \sum_{j=1}^K \log p_j(\mathbf{x}_i^{(x)}) \right)$	$f_k(\mathbf{x}) = \sum_{i=1}^N \left(\ln p_k(\mathbf{x}_i^{(x)}) - \frac{1}{K} \sum_{j=1}^K \ln p_j(\mathbf{x}_i^{(x)}) \right)$
2	196	2 行目	… $f(\mathbf{x}_i^{(x)})$ の評価に, …	… $f_k(\mathbf{x})$ の評価に, …
2	200	7 行目	$\partial \log \mathbf{X} / \partial \mathbf{X} = (\mathbf{X}^{-1})^T$ については, …	$\partial \ln \mathbf{X} / \partial \mathbf{X} = (\mathbf{X}^{-1})^T$ については, …
2	200	式(A.7)	$\frac{\partial \log \mathbf{X} }{\partial \mathbf{X}} = \frac{\partial \log \mathbf{X} }{\partial \mathbf{X} } \frac{\partial \mathbf{X} }{\partial \mathbf{X}} = \frac{1}{ \mathbf{X} } \frac{\partial \mathbf{X} }{\partial \mathbf{X}}$	$\frac{\partial \ln \mathbf{X} }{\partial \mathbf{X}} = \frac{\partial \ln \mathbf{X} }{\partial \mathbf{X} } \frac{\partial \mathbf{X} }{\partial \mathbf{X}} = \frac{1}{ \mathbf{X} } \frac{\partial \mathbf{X} }{\partial \mathbf{X}}$
2	200	式(A.10)	$\frac{\partial \log \mathbf{X} }{\partial \mathbf{X}} = \dots$	$\frac{\partial \ln \mathbf{X} }{\partial \mathbf{X}} = \dots$
5	201	第 2 章 2.2 4 行目	$E_D\{(y(\mathbf{x}; D) - E_D\{y(\mathbf{x}; D)\})\} = \dots$	$E_D\{(y(\mathbf{x}; D) - E_D\{y(\mathbf{x}; D)\})\} = \dots$
1	201	第 3 章 3.1(2)	(2)誤りの期待値: 式(3.11)の定義を和の形で表現すれば, $P_\varepsilon = \dots$	(2)ベイズ誤り率: 式(3.11)の定義を和の形で表現すれば, $\varepsilon^* = \dots$
1	202	5 行目	$P_\varepsilon =$	$\varepsilon^* = \dots$
2	203	下から 4 行目	$\sqrt{(x_1 - \mathbf{x})^2 + (y_1 - y)}$ が成り立っている…	$\sqrt{(x_1 - \mathbf{x})^2 + (y_1 - y)^2}$ が成り立っている…
1	204	8 行目	$= P(C_1 \mathbf{x}) P(C_2 \mathbf{x}) = r^*(\mathbf{x})(1 - r^*(\mathbf{x}))$	$= P(C_1 \mathbf{x}) P(C_2 \mathbf{x}) = \varepsilon(\mathbf{x})(1 - \varepsilon(\mathbf{x}))$
2	205	下から 3 行目	$= \frac{2}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \left(\mathbf{S}_B \mathbf{w} - \frac{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \mathbf{S}_W \mathbf{w} \right)$	$= \frac{2}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \left(\mathbf{S}_B \mathbf{w} - \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \mathbf{S}_W \mathbf{w} \right)$
5	206	解答 6.5 1 行目	環境 1 の失敗のオッズ比は $(0.01/0.99) \div (0.1/0.9) = 0.0091$,	環境 1 の失敗のオッズ比は $(0.01/0.99) \div (0.1/0.9) = 0.091$,

5	207	7.3 2行目	$\frac{dw_{ij}}{dt} = -\frac{\partial \tilde{E}(\mathbf{w})}{\partial w_{ij}} = \sum_{n=1}^N \delta_i^n V_j^n + \begin{cases} -\lambda & (\mathbf{w}_{ij} > 0 \text{のとき}) \\ +0 & (\mathbf{w}_{ij} = 0 \text{のとき}) \\ +\lambda & (\mathbf{w}_{ij} < 0 \text{のとき}) \end{cases}$	$\frac{dw_{kj}}{dt} = -\eta \frac{\partial \tilde{E}(\mathbf{w})}{\partial w_{kj}} = \eta \sum_{n=1}^N \delta_k^n V_j^n + \begin{cases} -\eta\lambda & (\mathbf{w}_{kj} > 0 \text{のとき}) \\ 0 & (\mathbf{w}_{kj} = 0 \text{のとき}) \\ +\eta\lambda & (\mathbf{w}_{kj} < 0 \text{のとき}) \end{cases}$
5	208	8.4 3行目	$= (1, u_1, u_1 u_2)(1, v_1, v_1 v_2)^T$	$= (1, u_1, u_2, u_1 u_2)(1, v_1, v_2, v_1 v_2)^T$
2	208	8.5(2) 2行目	$\left. \frac{\partial \tilde{L}}{\partial \mathbf{w}} \right _{\mathbf{w}=\mathbf{w}_0} = \dots$	$\left. \frac{\partial \tilde{L}_p}{\partial \mathbf{w}} \right _{\mathbf{w}=\mathbf{w}_0} = \dots$
2	208	8.5(2) 3行目	$\frac{\partial \tilde{L}}{\partial \xi_i} = \dots$	$\frac{\partial \tilde{L}_p}{\partial \xi_i} = \dots$
2	208	8.5(2) 4行目	$\frac{\partial \tilde{L}}{\partial \rho} = \dots$	$\frac{\partial \tilde{L}_p}{\partial \rho} = \dots$
2	209	9.3 1~3行目	$\mathbf{X}^T \mathbf{X} = (\mathbf{U} \Lambda \mathbf{V}^T)^T (\mathbf{U} \Lambda \mathbf{V}^T) = \mathbf{V} \Lambda \mathbf{U}^T \mathbf{U} \Lambda \mathbf{V}^T = \mathbf{V} \Lambda \Lambda \mathbf{V}^T$ となるので、 ... $= \mathbf{V} \Lambda \Lambda$ が成り立つ。したがって、 $(\mathbf{X}^T \mathbf{X} \mathbf{v}_1 \dots \mathbf{X}^T \mathbf{X} \mathbf{v}_d) = (\lambda_1 \mathbf{v}_1 \dots \lambda_d \mathbf{v}_d)$	$\mathbf{X}^T \mathbf{X} = (\mathbf{U} \Lambda \mathbf{V}^T)^T (\mathbf{U} \Lambda \mathbf{V}^T) = \mathbf{V} \Lambda^T \mathbf{U}^T \mathbf{U} \Lambda \mathbf{V}^T = \mathbf{V} \Lambda^T \Lambda \mathbf{V}^T$ となるので、... $= \mathbf{V} \Lambda^T \Lambda$ が成り立つ。したがって、 $(\mathbf{X}^T \mathbf{X} \mathbf{v}_1 \dots \mathbf{X}^T \mathbf{X} \mathbf{v}_p) = (\lambda_1 \mathbf{v}_1 \dots \lambda_p \mathbf{v}_p)$
5	210	10.4 6行目	$f\left(\sum_{i=1}^{M+1} \lambda_i x_i\right) \leq \lambda_{M+1} x_{M+1} + (1 - \lambda_{M+1}) f\left(\sum_{i=1}^M \alpha_i x_i\right)$	$f\left(\sum_{i=1}^{M+1} \lambda_i x_i\right) \leq \lambda_{M+1} f(x_{M+1}) + (1 - \lambda_{M+1}) f\left(\sum_{i=1}^M \alpha_i x_i\right)$
2	210	10.4 最下行	が成り立っていることは仮定したので、 $f\left(\sum_{i=1}^{M+1} \lambda_i x_i\right) \leq \sum_{i=1}^{M+1} \lambda_i f(x_i)$ が成り立つ。	が成り立っていることは仮定したので、 $f\left(\sum_{i=1}^{M+1} \lambda_i x_i\right) \leq \sum_{i=1}^{M+1} \lambda_i f(x_i)$ が成り立つ。
2	215	左段 11行目	AUR 33	AUC 330