# 深層学習 生成 AI の基礎

共立出版

最終更新：2024 年 4 月 6 日

## 正誤表

| ページ | 行数,<br>図・表・式番号 | 誤 | 正 |
|---|---|---|---|
| 17 | 3 行目 | $\ldots, 0 < l < L, z_j^{(0)} = x_j, \ldots$ | $\ldots, 1 < l < L, z_j^{(1)} = x_j, \ldots$ |
| 17 | 5 行目 | $\ldots, 0 < l < L, \mathbf{Z}^{(0)} = \mathbf{X}, \ldots$ | $\ldots, 1 < l < L, \mathbf{Z}^{(1)} = \mathbf{X}, \ldots$ |
| 17 | 下から 3 行目 | すべての重みを $\mathbf{W} = (\mathbf{W}^{(1)} \cdots \mathbf{W}^{(L)})$ とかき, | 「とかき,」に脚注として以下を追加.「行列 $\mathbf{W}^{(1)}$ はダミー.」 |
| 95 | 図 6.1 のキャプション | 図中の丸は現在の状態で，$\cdots$ | 図中の白丸は状態を表わし，$\cdots$ |
| 95 | 11 から 12 行目 | 図中の丸は現在の状態で，$\cdots$ | 図中の白丸は状態を表わし，$\cdots$ |
| 97 | 7 行目 | 一般に，強化学習では，学習のスタート時点では環境に対する情報はもっていないのが普通である． | 一般に，強化学習では，状態の集合と行動の集合は既知とするが，その他の環境に対する情報は，学習のスタート時点ではもっていないのが普通である． |
| 98 | 10 行目 | に依存する． | 「に依存する．」に脚注として以下を追加．「行動価値関数を，期待値の定義により書きくだすと $\begin{aligned} Q^\pi(s,a) &\equiv \mathbb{E}^\pi[C_0 \mid S_0 = s, A_0 = a] \\ &= g(s, a) + \sum_{k=1} \sum_{(s', a')} P(S_k = s', A_k = a') g(s', a') \\ &= g(s, a) + \sum_{k=1} \sum_{(s', a')} \pi(s' \mid a') P(S_k = s') g(s', a') \end{aligned}$ となる．方策 $\pi$ が定まっていても，一般に，$P(S_k = s')$ は確定しないので，行動価値関数は陽にはわからない．」 |
| 110 | 10-11 行目 | 確率の列（離散分布）であるから，その出力（の対数）を誤差とする誤差逆伝播で計算可能である． | 確率の列（離散分布）であり，$\ln \pi_{\boldsymbol{\theta}}(a_t \mid s_t)$ は出力の関数なので，それの出力ユニットの活性による微分を誤差とする誤差逆伝播で計算可能である． |
| 116 | 9 行目 | を考える． | 「を考える．」に脚注として以下を追加．「以下では，確率変数列のインデックスとして，下つきだけではなく，上つきもつかう．たとえば，文が複数あるときに，各文を区別するために下つきインデックスをもちいるので，列のインデックスとして上つきをつかう．」 |
| 116 | 下から 7 行目 | $\ldots, \mathbf{x}^{\mathrm{t}}$ などがでてくる． | $\ldots, \mathbf{x}^{t}$ などがでてくる． |

# 正誤表

| | | | |
|---|---|---|---|
| 123 | 下から 5 行目 式 (7.3.1) | $P(\mathbf{Y}\,\vert\,\mathbf{X}) = \prod_{t=1}^{T+1} P(y_t\,\vert\,\mathbf{Y}_{0:t-1}, \mathbf{X})$ | $P(\mathbf{Y}\,\vert\,\mathbf{X}) = \prod_{t=1}^{T+1} P(y^t\,\vert\,\mathbf{Y}^{0:t-1}, \mathbf{X})$ |
| 123 | 下から 4 行目 | ただし，$\mathbf{Y}_{0:t-1}$ は部分列 $y_0, y_1, \ldots, y_{t-1}$ を表わす. | ただし，$\mathbf{Y}^{0:t-1}$ は部分列 $y^0, y^1, \ldots, y^{t-1}$ を表わす. |
| 124 | 図 7.4 | 右上のほうの $\mathbf{c}_4$ | $\mathbf{c}_4 + \tilde{\mathbf{z}}_4$ |
| 125 | 図 7.5 (a), (b), (c) | 右上にある角がまるい四角の中にある $\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2$ はそれぞれ | $\mathbf{c}_0 + \tilde{\mathbf{z}}_0, \mathbf{c}_1 + \tilde{\mathbf{z}}_1, \mathbf{c}_2 + \tilde{\mathbf{z}}_2$ |
| 127, 128 | 図 7.6, 図 7.7 (a), (b), (c) | 右上にある角がまるい四角の中にある $\mathbf{c}_4, \mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2$ をそれぞれ | 削除 |
| 129 | 6 行目 | （ここでは 12）である. | 「（ここでは 12 である）.」に脚注として以下を追加.「自己注意の項で述べたように，関連性ベクトルの正規化でもちいられるソフトマックス関数の性質上，マルチヘッド化せずに一つのベクトルの注意を求めただけでは，正規化成分のうちの一つの成分の値が 1 に近く，その他の成分がほとんど 0 になることが起こる.」 |
| 220 | 6 行目 | $p(t_n = 0\,\vert\,\mathbf{x}_n)$ $= (1-\theta_1)\cdot p(t_n^* = 0\,\vert\,\mathbf{x}_n) + \theta_1 \cdot p(t_n^* = 1\,\vert\,\mathbf{x}_n),$ | $p(t_n = 0\,\vert\,\mathbf{x}_n)$ $= (1-\theta_0)\cdot p(t_n^* = 0\,\vert\,\mathbf{x}_n) + \theta_1 \cdot p(t_n^* = 1\,\vert\,\mathbf{x}_n),$ |
| 220 | 7 行目 | $p(t_n = 1\,\vert\,\mathbf{x}_n)$ $= \theta_0 \cdot p(t_n^* = 0\,\vert\,\mathbf{x}_n) + (1-\theta_0)\cdot p(t_n^* = 1\,\vert\,\mathbf{x}_n)$ | $p(t_n = 1\,\vert\,\mathbf{x}_n)$ $= \theta_0 \cdot p(t_n^* = 0\,\vert\,\mathbf{x}_n) + (1-\theta_1)\cdot p(t_n^* = 1\,\vert\,\mathbf{x}_n)$ |
| 220 | 12, 13 行目 | $\begin{cases} p(t_n = 0\,\vert\,\mathbf{x}_n) = (1-\theta_1)\cdot(1-y_n) + \theta_1 \cdot y_n, \\ p(t_n = 1\,\vert\,\mathbf{x}_n) = \theta_0 \cdot (1-y_n) + (1-\theta_0)\cdot y_n \end{cases}$ | $\begin{cases} p(t_n = 0\,\vert\,\mathbf{x}_n) = (1-\theta_0)\cdot(1-y_n) + \theta_1 \cdot y_n, \\ p(t_n = 1\,\vert\,\mathbf{x}_n) = \theta_0 \cdot (1-y_n) + (1-\theta_1)\cdot y_n \end{cases}$ |
| 226 | 4 行目と 12 行目 | こえては近づかない. | こえては離れない. |