Errata for the paper entitled

“Second-order stagewise backpropagation for Hessian-matrix analyses and investigation of negative curvature”

by Eiji Mizutani and Stuart E. Dreyfus.

— 2008 Special Issue: Advances in Neural Networks Research: IJCNN’07.

As of July 27, 2008.

(1) Page 194: In Eq. (4), \( \theta^s \) should be \( \theta^{s,s+1} \), and add “=” (notation for being equal to) between \( G^s \) and \( y^s_s + \delta^{s+1} \).

(2) Page 195: Eq. (5) should be expressed for recursion as

\[
Z^s = \frac{N^s,s+1}{P_s \times P_s} Z^{s+1} + \frac{N^s,s+1}{P_{s+1} \times P_{s+1}} \frac{\partial y^s_s}{\partial x^s} \xi^s,
\]

where all the notations are defined in the paper except a matrix \( N \) below

\[
N^s,s+1 = \def \frac{\partial x^{s+1}}{\partial x^s},
\]

The above recurrence relation is the same as Eq. (13) in Mizutani, Dreyfus, & Demmel (2005); see also Eq. (27) in Mizutani & Dreyfus, 2006.

(3) Page 195: In Eqs. (10) and (11), \( \theta^s \) should be \( \theta^{s,s+1} \); likewise, \( \theta^r \) should be \( \theta^{r,r+1} \).

(4) Page 197: Just above Eq. (14), the sentence

In general, the residual Hessian matrix \( S \) in \( H \) can separate into two types of blocks

should read

In general, the residual Hessian matrix \( S \) in \( H \) always includes two types of blocks

(5) Page 197: Eq. (16) should have included an additional matrix \( T \) for a general expression as

\[
S = \begin{bmatrix}
\sum_{n_3 \times n_3}^V & \sum_{n_2 \times n_2}^V & \sum_{n_1 \times n_1}^V \\
\sum_{n_2 \times n_2}^V & \sum_{n_1 \times n_1}^V & \sum_{n_1 \times n_1}^V \\
\sum_{n_1 \times n_1}^V & \sum_{n_1 \times n_1}^V & \sum_{n_1 \times n_1}^V
\end{bmatrix} + 
\begin{bmatrix}
\Gamma^{1,1T} & \Gamma^{1,2T} \\
\Gamma^{2,1T} & \Gamma^{2,2T} \\
\Gamma^{3,1T} & \Gamma^{3,2T}
\end{bmatrix}
+ T,
\]

where \( T \) denotes all the remaining terms; see Eqs. (22) and (23) in Mizutani, Dreyfus, & Demmel (2005) for such terms that construct \( T \). See also Mizutani 2008.

(6) Page 201: In Section 5, the second sentence

In MLP-learning, special sparsity structure inevitably arises in \( S \),

should read

In single-hidden-layer linear-output MLP-learning, special sparsity structure inevitably arises in \( S \) (e.g., see \( S \) in Proof of Lemma 2),
References:


Author information:

- **Eiji Mizutani** (e-mail: eiji@mail.ntust.edu.tw) is with Department of Industrial Management, National Taiwan University of Science & Technology, Taipei, TAIWAN.

- **Stuart Dreyfus** (e-mail: dreyfus@ieor.berkeley.edu) is with Department of Industrial Engineering & Operations Research, University of California, Berkeley, CA USA.