Author Correction for

TensorProjection Layer: A Tensor-Based Dimension Reduction Method in Deep Neural Networks

Toshinari Morimoto Su-Yun Huang

October 2025

Correction for Equation (11)

The authors found that in the published paper, the derivative

$$\frac{\partial \operatorname{vec}(G_k)}{\partial \operatorname{vec}(M_k)}, \text{ where } G_k = M_k^{-1/2},$$

was not correct. The corrected expression is given by (1). This correction does **not** affect the numerical results, since all experiments used automatic differentiation and did not rely on the analytical backpropagation formula.

Let

$$H_k = M_k^{1/2} \quad \Rightarrow \quad G_k = H_k^{-1},$$

where each $G_k, H_k, M_k \in \mathbb{R}^{q_k \times q_k}$ and M_k is symmetric positive definite.

We aim to find the Jacobian via the chain rule:

$$\frac{\partial \operatorname{vec}(G_k)}{\partial \operatorname{vec}(M_k)^\top} = \frac{\partial \operatorname{vec}(G_k)}{\partial \operatorname{vec}(H_k)^\top} \frac{\partial \operatorname{vec}(H_k)}{\partial \operatorname{vec}(M_k)^\top}.$$

—

Step 1: Relation between G_k and H_k

From $G_k H_k = I_{q_k}$, we have

$$G_k(\Delta H_k) + (\Delta G_k) H_k = O.$$

Then

$$\Delta G_k = -\,G_k(\Delta H_k) H_k^{-1} = -\,G_k(\Delta H_k) G_k.$$

Vectorizing both sides gives

$$\operatorname{vec}(\Delta G_k) = -(H_k^\top \otimes G_k) \operatorname{vec}(\Delta H_k),$$

so that
$$\frac{\partial\,\mathrm{vec}(G_k)}{\partial\,\mathrm{vec}(H_k)^\top} = -(H_k^\top \otimes G_k) = -(H_k^{-1} \otimes H_k^{-1}) = -(M_k^{-1/2} \otimes M_k^{-1/2}). \tag{1}$$

—

Step 2: Relation between $\boldsymbol{H_k}$ and $\boldsymbol{M_k}$

Since $H_k^2 = M_k$, we have
$$\Delta M_k = H_k(\Delta H_k) + (\Delta H_k)H_k.$$

Using the identity $\mathrm{vec}(AXB) = (B^\top \otimes A)\,\mathrm{vec}(X)$,

$$\begin{aligned}\mathrm{vec}(\Delta M_k) &= (I_{q_k} \otimes H_k)\,\mathrm{vec}(\Delta H_k) + (H_k \otimes I_{q_k})\,\mathrm{vec}(\Delta H_k)\\ &= \big(I_{q_k} \otimes H_k + H_k \otimes I_{q_k}\big)\,\mathrm{vec}(\Delta H_k).\end{aligned}$$

Hence,
$$\frac{\partial\,\mathrm{vec}(H_k)}{\partial\,\mathrm{vec}(M_k)^\top} = \big(I_{q_k} \otimes H_k + H_k \otimes I_{q_k}\big)^{-1}. \tag{2}$$

—

Step 3: Chain rule

Combining (1) and (2),

$$\begin{aligned}\frac{\partial\,\mathrm{vec}(G_k)}{\partial\,\mathrm{vec}(M_k)^\top} &= \frac{\partial\,\mathrm{vec}(G_k)}{\partial\,\mathrm{vec}(H_k)^\top}\frac{\partial\,\mathrm{vec}(H_k)}{\partial\,\mathrm{vec}(M_k)^\top}\\ &= -(H_k^{-1} \otimes H_k^{-1})\big(I_{q_k} \otimes H_k + H_k \otimes I_{q_k}\big)^{-1}.\end{aligned}$$

Substituting $H_k = M_k^{1/2}$ gives the corrected result:

$$\boxed{\frac{\partial\,\mathrm{vec}(G_k)}{\partial\,\mathrm{vec}(M_k)^\top} = -(M_k^{-1/2} \otimes M_k^{-1/2})\big(I_{q_k} \otimes M_k^{1/2} + M_k^{1/2} \otimes I_{q_k}\big)^{-1}.} \tag{1}$$

—

Invertibility remark

The matrix
$$I_{q_k} \otimes M_k^{1/2} + M_k^{1/2} \otimes I_{q_k}$$
is invertible if and only if M_k is positive definite, since its eigenvalues are $\{\sqrt{\lambda_i} + \sqrt{\lambda_j}\}_{i,j}$, where λ_i are the eigenvalues of M_k.