

Ograniczone maszyny Boltzmanna dla danych o wartościach zespolonych

Maciej Jaworski

Politechnika Krakowska

e-mail: maciej.jaworski@pk.edu.pl

Popularne ostatnimi czasy modele generatywne, takie jak GAN czy Stable Diffusion, osiągają wysoką jakość generowanych danych, lecz są trudne w trenowaniu i często brakuje im interpretowalności. W przeciwieństwie do nich, Ograniczone Maszyny Boltzmanna (RBM) [4], choć o mniejszych możliwościach, oferują prostą i dobrze rozumianą strukturę probabilistyczną. Modelują bezpośrednio rozkład Boltzmanna możliwych stanów sieci, a ich trening oparty jest na minimalizacji funkcji energii.

W pracy prezentujemy rozszerzenie sieci RBM do przestrzeni zespolonej (complex-valued RBM, cvRBM) [2], w której wejścia, wagi i biasy warstwy wejściowej są liczbami zespolonymi.

Model cvRBM

Standardowa energia w sieci RBM dana jest wzorem:

$$E(v, h) = -v^T W h - b^T v - c^T h,$$

gdzie v i h to odpowiednio warstwy widzialna i ukryta sieci, W to macierz wag łączących neurony obu warstw, a b i c to wektory biasów dla warstw v i h . W wersji zespolonej, by energia pozostała rzeczywista, definiujemy:

$$E(v, h) = -\Re(v^\dagger W h) - \Re(b^\dagger v) - c^T h$$

Przy oznaczeniach:

$$v = \Re(v) + i\Im(v), \quad W = \Re(W) + i\Im(W), \quad b = \Re(b) + i\Im(b)$$

otrzymujemy:

$$E(v, h) = -\Re(v)^T \Re(W) h - \Im(v)^T \Im(W) h - \Re(b)^T \Re(v) - \Im(b)^T \Im(v) - c^T h$$

czyli model ekwiwalentny do dwóch połączonych sieci RBM z dzieleniem informacji na komponent rzeczywisty i urojony.

Eksperymenty

Badania przeprowadzono na zbiorze MNIST. W wersji klasycznej użyto oryginalnych obrazów, natomiast w cvRBM dane wejściowe to $v = x + iy$, gdzie x to obraz, a y to wynik filtru Sobela.

Do uczenia zastosowano algorytm Contrastive Divergence (CD-1) [1], 300 epok, różne liczby neuronów ukrytych (20-220). Porównano błąd rekonstrukcji i wygenerowane próbki. Wyniki wstępne pokazują, że cvRBM uchwytuje osiąga dokładności porównywalne z klasyczną siecią RBM.

Dalsze kierunki badań

Badania wykonane w ramach pracy mogą zostać poszerzone w wielu interesujących kierunkach:

- alternatywne funkcje energii (np. z pełnym sprzężeniem hermitowskim),
- zastosowanie do innych danych (fizycznych, kwantowych),
- rozszerzenie do gaussowskich sieci cvRBM [3].

MSC (2020): 62M45, 68T07

Literatura

- [1] G. Hinton, *Training Products of Experts by Minimizing Contrastive Divergence*, Neural Computation, 2022, 14 1771–1800.
- [2] M. Jaworski, A. Niemczynowicz, R. Kycia, *Complex-valued Restricted Boltzmann Machines for Learning Image Data Distributions*, Proc. of the International Conference on Artificial Intelligence and Soft Computing, 2024.
- [3] T. Nakashika, S. Takaki, J. Yamagishi, *Complex-Valued Restricted Boltzmann Machine for Direct Speech Parameterization from Complex Spectra*, Proc. of the Interspeech 2017, 4021–4025.
- [4] P. Smolensky, *Chapter 6: Information Processing in Dynamical Systems: Foundations of Harmony Theory*, Rumelhart, David E.; McClelland, James L. (eds.). Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations, MIT Press, 1986, 194–281.