Code Trick For TransG

February 21, 2017

1. The class "TransG" is the experimental version, rather than "TransG_Hierarchical". Note that, for numerical stabilization, we fix the variance σ as a constant.

2. When generating the new cluster, we assign the CRP factor as the mixture factor, which is required by CRP theoretically. But as to the center of new cluster, we assign a random vector rather than $\mathbf{t} - \mathbf{h}$. Because in big data scenario, theoretical center is far away from the ground-truth.

3. Regarding the learning methodology of parameter π and σ, we applied the stochastic gradient ascent (SGD) for efficiency, rather than likelihood counting, which is naturally suitable for CRP but inefficient. We suggest the reader to implement both for experiments.