Theoretical background

The proposed problem is the problem of stochastic matrix factorization. For it to have a unique solution, it has to be regularized as follows:

$$\sum_{d,w} n_{dw} \ln \sum_{t} \phi_{wt}\theta_{td} + R(\Phi, \Theta) \to \max_{\Phi,\Theta} \tag{1}$$

Where $R(\Phi, \Theta)$ — additive regularization term. This is the classic problem of ARTM. BigARTM authors proposed a solution using expectation-maximization algorithm:

$$\begin{cases} p_{tdw} \equiv p(t|d,w) = \underset{t \in T}{norm} \\ \phi_{wt} = \underset{w \in W}{norm} \left(n_{wt} + \phi_{wt} \frac{\partial R}{\partial \phi_{wt}} \right), \quad n_{wt} = \sum\limits_{d \in D} n_{dw} p_{tdw} \\ \theta_{td} = \underset{t \in T}{norm} \left(n_{td} + \theta_{td} \frac{\partial R}{\partial \theta_{td}} \right), \quad n_{td} = \sum\limits_{w \in d} n_{dw} p_{tdw} \end{cases} \tag{2}$$

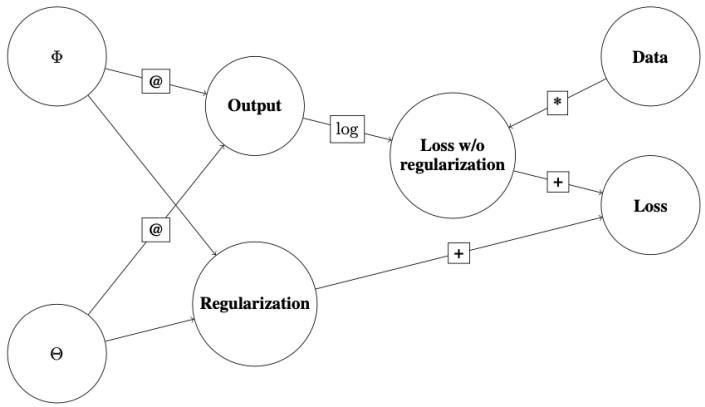

Figure 1: Calculation graph