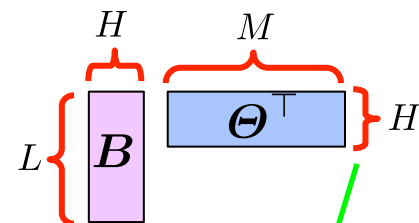
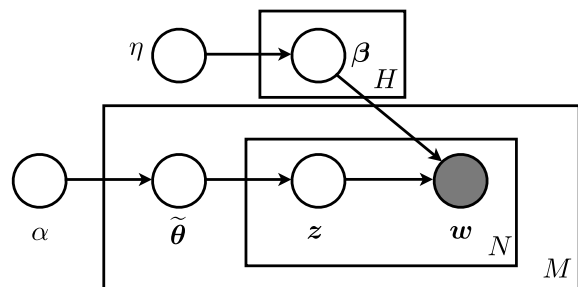


T-30: 変分ベイズLDAの漸近解析

中島伸一 (ベルリン工大)、佐藤一誠 (東大)、杉山将 (東大)
渡辺一帆 (豊橋技大)、小林寛子 (ニコン)

Latent Dirichlet allocation (LDA)



$$p(\{\mathbf{w}^{(n,m)}\} | \boldsymbol{\Theta}, \mathbf{B}) = \prod_{m=1}^M \prod_{n=1}^{N^{(m)}} \prod_{l=1}^L \left((B\boldsymbol{\Theta}^\top)_{l,m} \right)^{w_l^{(n,m)}}$$

- ドキュメントあたりの単語数 N 、ドキュメント数 M 、語彙サイズ L が大きい極限での自由エネルギーを解析。
- $N, M \gg L$ のときのスパース性を理論的に解明。最尤法との比較。

$$F = S + \lambda \log N + o_P(\cdot)$$

