Collapsed Inference Algorithms (cont'd.)

$$P(\boldsymbol{W}, \boldsymbol{Z} \mid \alpha, \beta) = \prod_{j=1}^{M} \frac{B(\alpha + \sigma_j)}{B(\alpha)} \prod_{i=1}^{K} \frac{B(\beta + \delta_i)}{B(\beta)}$$

- Distribution over only **W** and **Z**
- How do we get back Θ and Φ? MAP estimate from **Z**:

$\widehat{\theta_{j,k}} = \frac{\sigma_{j,k} + \alpha_k}{\sum_{i=1}^{K} \sigma_{j,i} + \alpha_i}$ and $\widehat{\phi_{k,v}} = \frac{\delta_{k,v} + \beta_v}{\sum_{r=1}^{V} \delta_{k,r} + \beta_r}$

- But how do we get the count vectors σ_j and δ_k?
  1. Gibbs sampling: sample values of **Z**, count from those samples
  2. Variational inference: use a surrogate distribution to model the probability of **Z** and use its expectation ("expected counts")