Maximum a Posteriori (MAP) Estimate

$$\Lambda^* = \arg\max_{\Lambda} p(\Lambda)\, p(Data \mid \Lambda)$$

- We may use p(Λ) to encode all kinds of preferences and constraints, e.g.,
  - p(Λ)>0 if and only if one topic is precisely "background": p(w|θ_B)
  - p(Λ)>0 if and only if for a particular doc d, $\pi_{d,3}$=0 and $\pi_{d,1}$=1/2
  - p(Λ) favors a Λ with topics that assign high probabilities to some particular words
- The MAP estimate (with conjugate prior) can be computed using a similar EM algorithm to the ML estimate with smoothing to reflect prior preferences