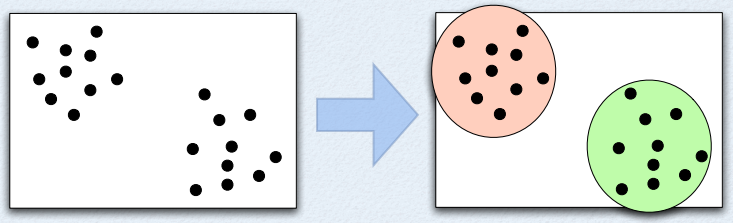
Accelerated Variational Dirichlet Process Mixture Models

東京工業大学 計算工学専攻
佐藤研究室 栗原賢一

clustering

- to make clusters so each cluster has similar data.



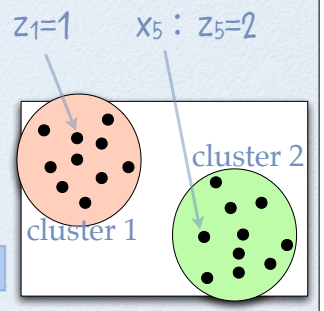
probabilistic model: mixture model

e.g. mixture of Gaussians

- generative model

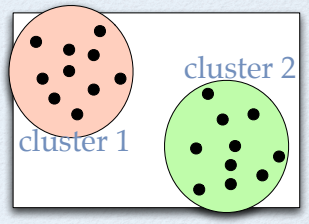
1. choose cluster z_n
   - $z_n \sim p(z_n \mid \pi) \equiv \pi_{z_n}$
     - note $\pi_1 + \pi_2 = 1$
2. sample x_n from cluster z_n
   - $x_n \sim p(x_n \mid \eta_{z_n})$

$\eta_{z_n} = (\mu_{z_n}, \Sigma_{z_n})$
mean variance



inference of assignments

- $p(Z|X)$
- e.g. $p(z_n=1|x_n) \lessgtr p(z_n=2|x_n)$



inference of #clusters

- posterior $p(K|D)$
  - $p(K|D) \propto p(K,D) = p(K)\, p(D|K)$

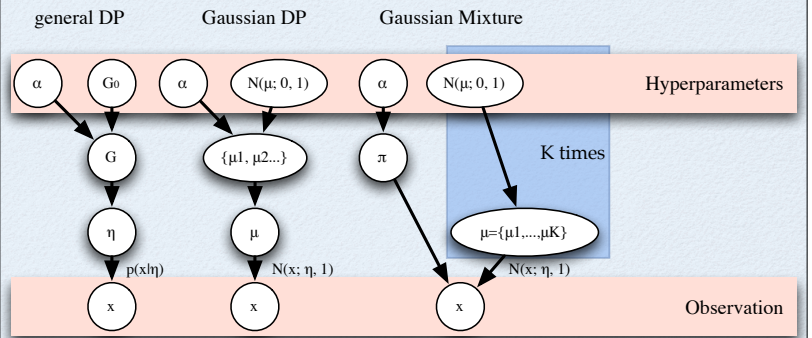
what is $p(K)$???

Dirichlet process

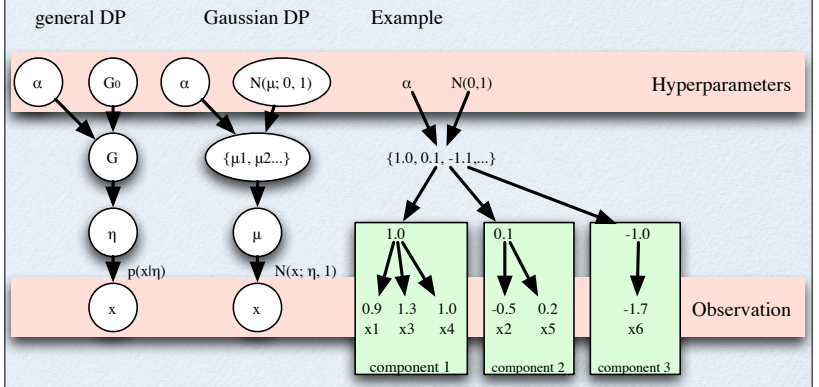
Dirichlet process

- #components = ∞
  - i.e. $K = \infty$ $\therefore$ $p(K)$ is not required
  - we observe " finite" components.

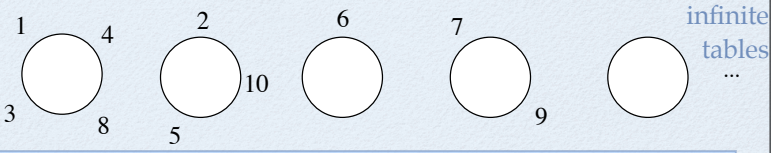
Finite and Infinite Mixtures



G0 and an example



Chinese restaurant process



4/10+α 3/10+α 1/10+α 2/10+α α/10+α

prob. of next assignment

$$p(z_N|[z_1 \ldots z_{N-1}]) = \begin{cases} \frac{N_c^{-N}}{\alpha+N-1} & (N_c^{-N} > 0; \\ & Z_N \text{ is an existing cluster.}) \\ \frac{\alpha}{\alpha+N-1} & (N_c^{-N} = 0; \\ & Z_N \text{ is a new cluster.}) \end{cases}$$

discussion: p(Z, π, K), p([Z])

- non-DP p(Z, π | K)

$$p(Z|\pi, K) = \prod_{c=1}^{K} \pi_c^{N_c}$$

$$D(\pi|K) = \frac{\Gamma(\alpha)}{\Gamma(\alpha/K)^K} \prod_{i=1}^{K} \pi_i^{\alpha/K-1}$$

$$p(K)$$

integral out π & take the limit of K

- DP p([Z])

$$p([Z]) = \alpha^{K_+} \left(\prod_{k=1}^{K_+} \prod_{j=1}^{N_k-1} j \right) \frac{\Gamma(\alpha)}{\Gamma(N+\alpha)}$$

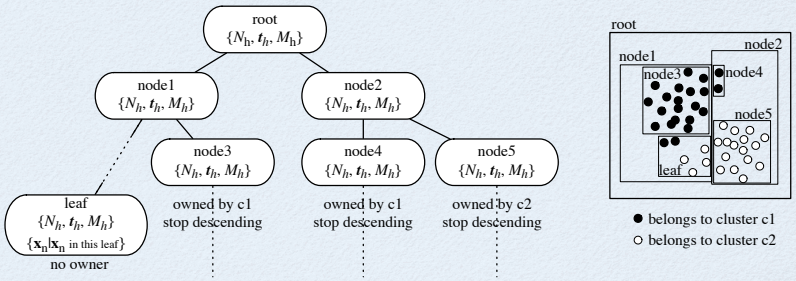
"DP is an option of p(Z, π, K)"

Accelerated Variational DP

points

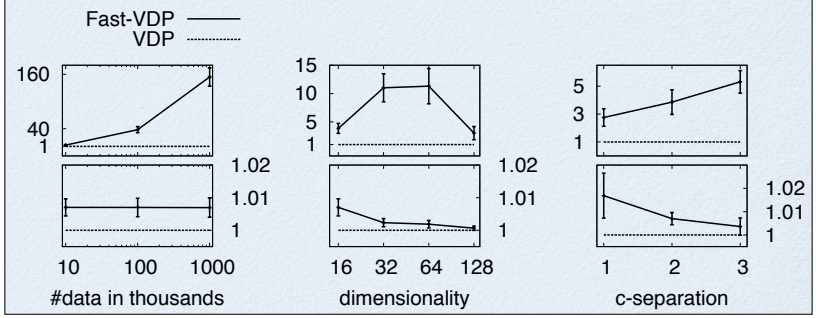
- variational DP
  - + acceleration by kd-trees
  - + relaxed truncation

acceleration by cache



Accelerated VDP

- 100 times faster than naive VDP on million data cases



relaxed truncation

- Blei's VDP
  - finite!
  - $\text{model}_T \not\subset \text{model}_{T+1}$
- Our VDP
  - infinite
  - $\text{model}_T \subset \text{model}_{T+1}$

thank you :)