Weekly Report IV

Donglai Wei

2010.5.11

0)Understanding Concentration parameter

I)DPmixture

Concentration parameter: α
$\mathcal{F}([z]) = \sum_{c=1}^{K}[\frac{DN_c}{2}log\pi + \frac{D}{2}log\frac{\xi_c}{\xi_0}logdet(B_c) - \frac{\eta_0}{2}logdet(B_0) - log\frac{\Gamma_D(\frac{\eta_c}{2})}{\Gamma_D(\frac{\eta_0}{2})}\ \underline{-log(\Gamma(N_c)) - log\alpha] + log\frac{\Gamma(N+\alpha)}{\Gamma(\alpha)}}$

a)

Given Ground Truth of z^*, $\mathcal{F}([z^*])$ is a function of α:
$\frac{\partial\mathcal{F}([z^*])}{\partial\alpha} = \Psi(N+\alpha) - \Psi(\alpha) - \dfrac{K}{\alpha}$

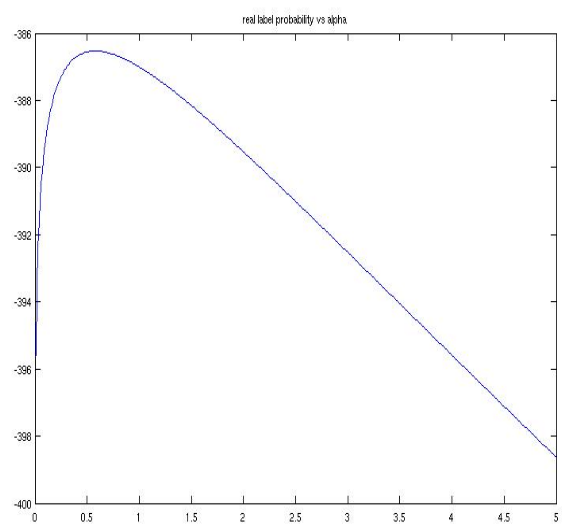


Figure 1: domain:0-5

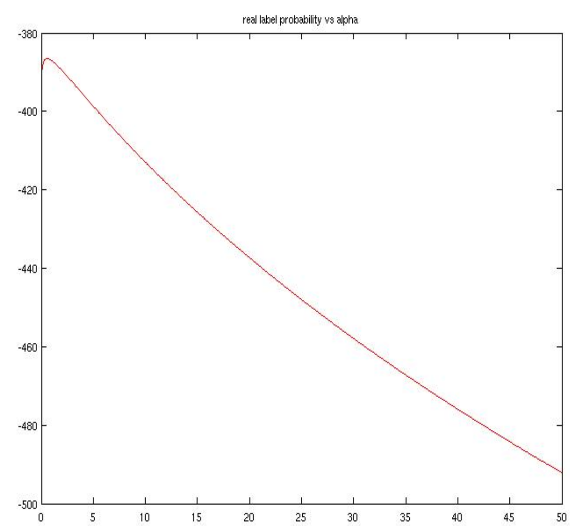


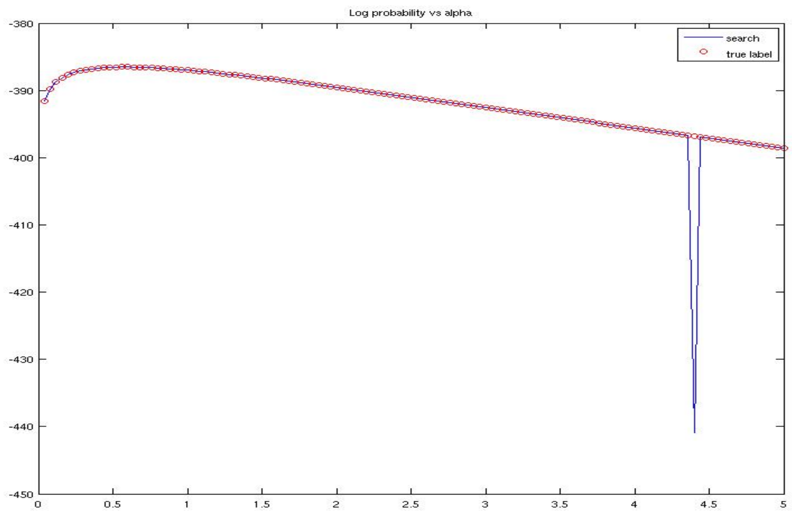
Figure 2: domain:0-50

b)

Given $\alpha, \mathcal{F}([z])$
$= (Likelihood\ term) \sum_{c=1}^{K}[\frac{DN_c}{2}log\pi + \frac{D}{2}log\frac{\xi_c}{\xi_0}logdet(B_c) - \frac{\eta_0}{2}logdet(B_0) - log\frac{\Gamma_D(\frac{\eta_c}{2})}{\Gamma_D(\frac{\eta_0}{2})}]$

$(Allocation\ term) - \sum_{c=1}^{K}[log(\Gamma(N_c)) + log\alpha]$

$(Constant\ term) + log\frac{\Gamma(N+\alpha)}{\Gamma(\alpha)}$



II) HDP(multinomial)

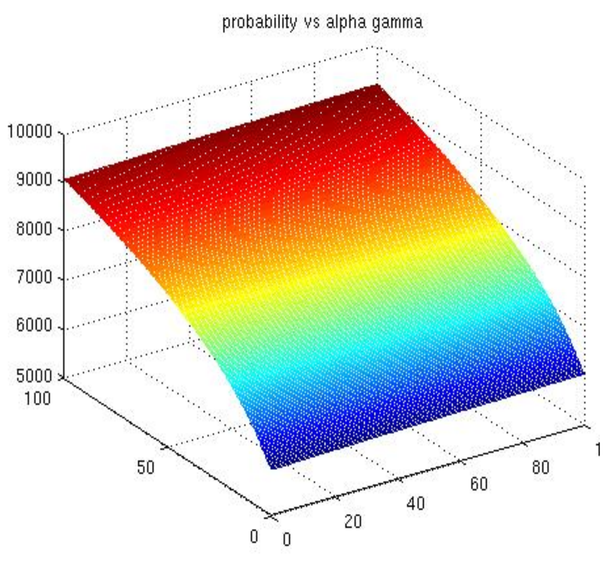


Figure 3: Log-probability

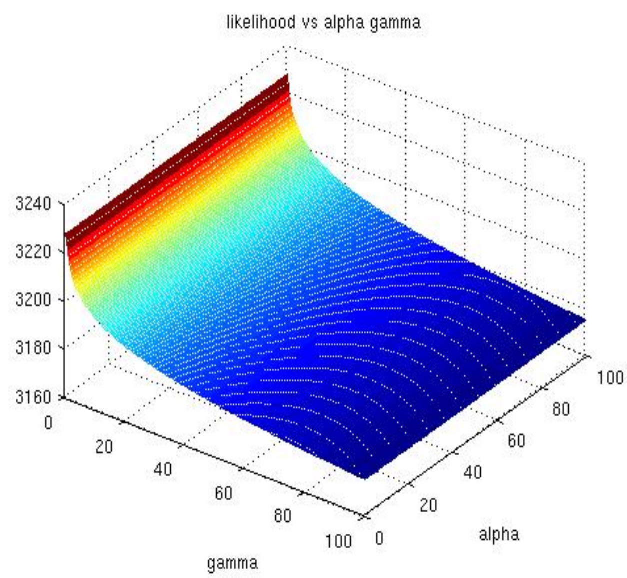


Figure 4: Likelihood term

1)Normal-Inverse Wishart

$-logp(x|z, \lambda)$
$=$

(Likelihood)$\sum_{k=1}^{K}[\frac{Dn_{..k}}{2}log\pi + \frac{D}{2}log\frac{\xi_k}{\xi_0} + \frac{\eta_k}{2}logdet(B_k) - \frac{\eta_0}{2}logdet(B_0) - log\frac{\Gamma_D(\frac{\eta_k}{2})}{\Gamma_D(\frac{\eta_0}{2})}]$
$+$
(Allocation:)$\underline{\sum_{j=1}^{J}\sum_{t=1}^{m_{j.}}[\frac{1}{m_{j.}}log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)} - log(\Gamma(n_{jt.}) - log\alpha] + \sum_{k=1}^{K}[\frac{1}{K}log\frac{\Gamma(T+\gamma)}{\Gamma(\gamma)} - log(\Gamma(m_{.k}) - log\gamma]}$

$=$
(t-term)$\sum_{j=1}^{J}\sum_{t=1}^{m_{j.}}\underline{[\frac{1}{m_{j.}}log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)} - log(\Gamma(n_{jt.}) - log\alpha + \frac{1}{Jm_{j.}}log\frac{\Gamma(T+\gamma)}{\Gamma(\gamma)}]}$

+(k-term)$\sum_{k=1}^{K}[\frac{n_{..k}D}{2}log\pi + \frac{D}{2}log\frac{\xi_k}{\xi_0} + \frac{\eta_k}{2}logdet(B_k) - \frac{\eta_0}{2}logdet(B_0) - log\frac{\Gamma_D(\frac{\eta_k}{2})}{\Gamma_D(\frac{\eta_0}{2})} - \underline{log(\Gamma(m_{.k}) - log\gamma]}$

$=$
(DP mixture of z_{ji})$\sum_{k=1}^{K}[\frac{n_{..k}D}{2}log\pi + \frac{D}{2}log\frac{\xi_k}{\xi_0} + \frac{\eta_k}{2}logdet(B_k) - \frac{\eta_0}{2}logdet(B_0) - log\frac{\Gamma_D(\frac{\eta_k}{2})}{\Gamma_D(\frac{\eta_0}{2})} - log(\Gamma(n_{..k})) - log\gamma]$

-(m term)$\sum_{k=1}^{K}log[\dfrac{\Pi_{j=1}^{J}(\Gamma(n_{j.k})}{\Gamma(n_{..k})}] - log\dfrac{\Pi_{k=1}^{K}\Gamma(m_{.k})}{\Gamma(T+\gamma)} - Tlog(\alpha)$

+(constant)$\sum_{j=1}^{J}log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)} + log\Gamma(\gamma)$

where:

$\xi_k = \xi_0 + n_{..k}$
$m_k = \frac{n_{..k}\vec{x}_{(k_{jt_{ji}}=k)}+\xi_0 m_0}{\xi_k}$
$\eta_k = \eta_0 = n_{..k}$
$B_k = B_0 + n_{..k}S_k + \frac{n_{..k}\xi_0}{\xi_k}(\vec{x}_{(k_{jt_{ji}}=k)} - m_0)(\vec{x}_{(k_{jt_{ji}}=k)} - m_0)^T$
$S_k : samplecovariance$

2)Dirichlet-Multinomial

W:number of unique words
$n_{..k}^{w}$number of occurence of word w in dish k

$-logp(x|z,\lambda)$
$=$
(Likelihood)$\sum_{k=1}^{K}[log(\frac{\Gamma(n_{..k}+W\phi_0)}{\Gamma(W\phi_0)}) + \sum_{w=1}^{W}log(\frac{\Gamma(\phi_0)}{\Gamma(\phi_0+n_{..k}^{w})})]$
$+$
(Allocation:)$\underline{\sum_{j=1}^{J}\sum_{t=1}^{m_{j.}}[\frac{1}{m_{j.}}log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)} - log(\Gamma(n_{jt.}) - log\alpha] + \sum_{k=1}^{K}[\frac{1}{K}log\frac{\Gamma(T+\gamma)}{\Gamma(\gamma)} - log(\Gamma(m_{.k}) - log\gamma]}$

$=$
(t-term)$\underline{log\frac{\Gamma(T+\gamma)}{\Gamma(\gamma)} + \sum_{j=1}^{J}\{log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)} - \sum_{t=1}^{m_{j.}}[log(\Gamma(n_{jt.}) + log\alpha]\}}$

+(k-term)$\sum_{k=1}^{K}[log(\frac{\Gamma(n_{..k}+W\phi_0)}{\Gamma(W\phi_0)}) + log(\Pi_{w=1}^{W}\frac{\Gamma(\phi_0)}{\Gamma(\phi_0+n_{..k}^{w})}) - \underline{log(\Gamma(m_{.k}) - log\gamma]}$

$=$
(k-term)+$\sum_{k=1}^{K}[log(\frac{\Gamma(n_{..k}+W\phi_0)}{\Pi_1^W\Gamma(\phi_0+n_{..k}^{w})\Gamma(m_{.k})\Pi_{q=1}^{m_{.k}}\Gamma(n_{q.})})]$
(hyper-term)$-Tlog\alpha - Klog\gamma + Klog\frac{\Gamma(\phi_0)^W}{\Gamma(W\phi_0)} + log\Gamma(T+\gamma)$
(constant-term)$-log\Gamma(\gamma) + \sum_{j=1}^{J}log\frac{\Gamma(n_{j..}+\alpha)}{\Gamma(\alpha)}$

Experiment

$\alpha = 0.5, \gamma = 1.5$

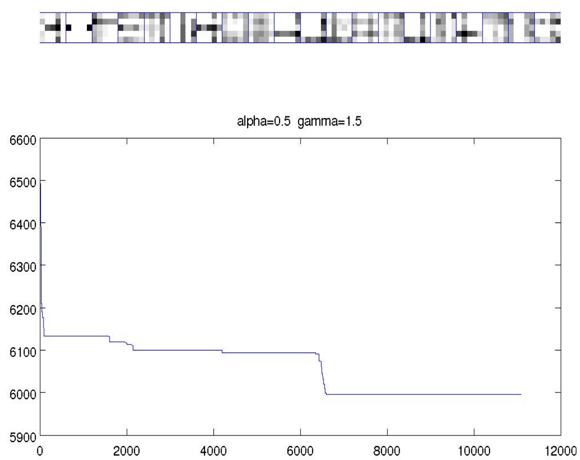


Figure 5: ME result

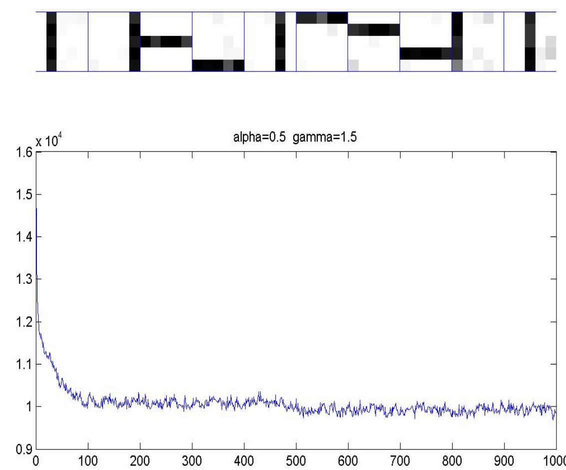


Figure 6: Gibbs

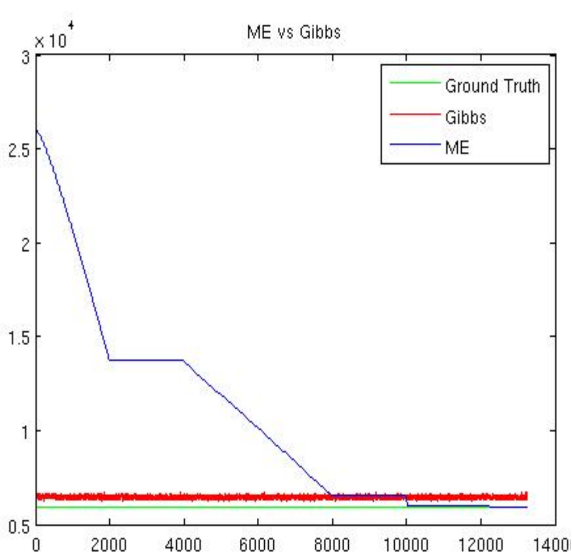


Figure 7: ME vs Gibbs

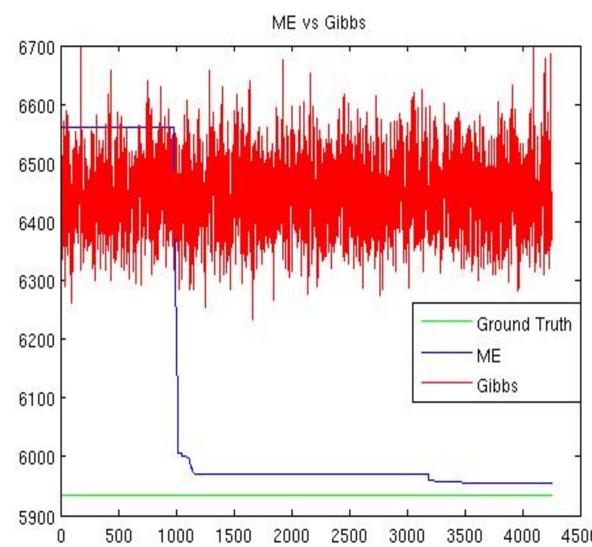


Figure 8: ME vs Gibbs