Finaally...Let there be bars

Donglai Wei

2010.5.25

1) Outline

(I) Initialization:

  (1) Naive(bottom-up,top-down)

  (2) Auxilary Dishes

    (A) Bottom-up:undesirable for large data setcounter

    (B) **Top-down:**

      (a) split-table:Bad global maxima (Since auxilary dishes can be uninformative)

      (b) approximation:**greedy+drop t-term**

(II) Search

  (1) Stochastic search

  (2) Big Moves

1.1) How does the restaurants look like

2) Initialization

2.1) Split-table Can be BAD

Below, I make 100 runs of split-table for the initialization(Auxilary Dishes) of Restaurant 1.

Later, I checked that assign the whole restaurant to the second Auxilary Dish(Restaurant) is the global maxima.

One rescue is to prune the Auxilary Dishes beforehand...But it can be tricky....

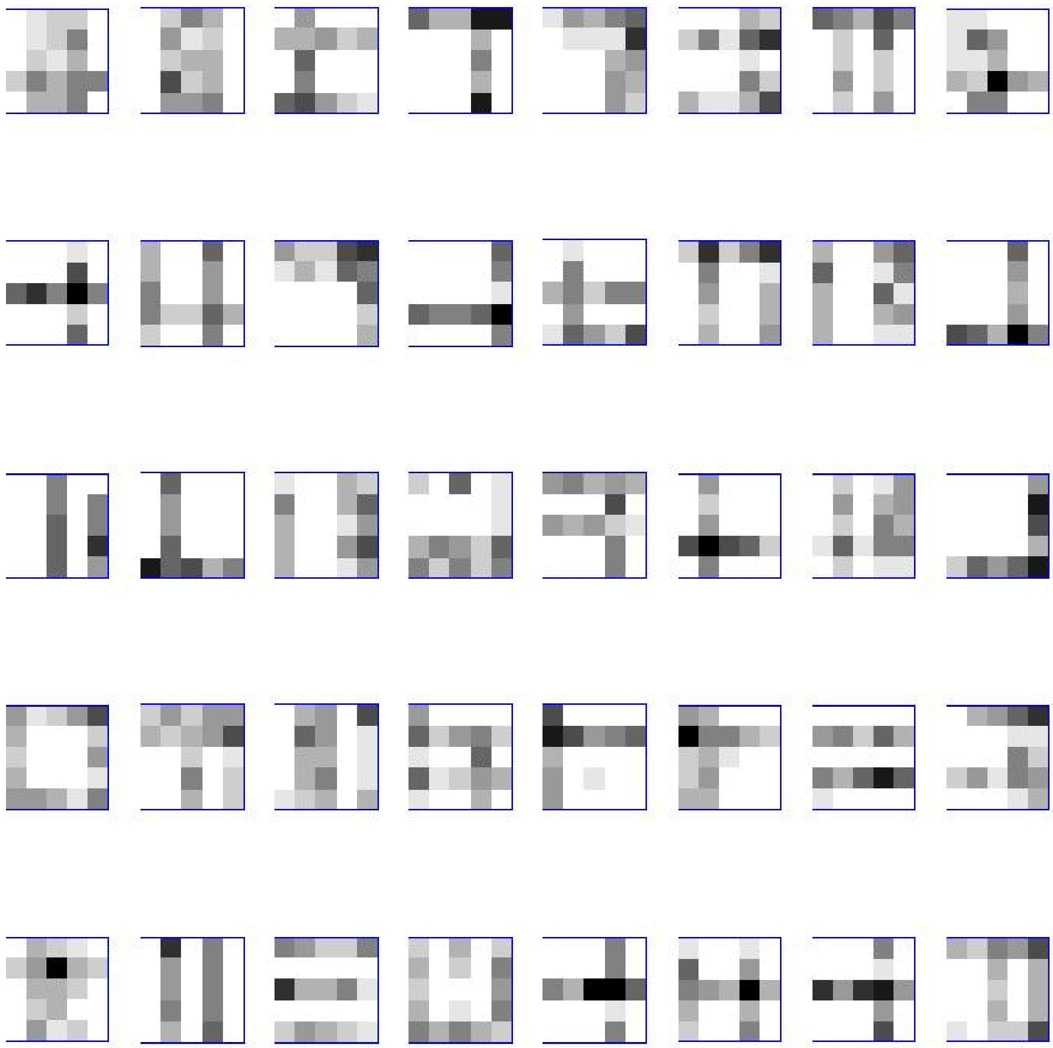

Figure 1: 40 Restaurants

Figure 2: Initialization For the first Restaurant

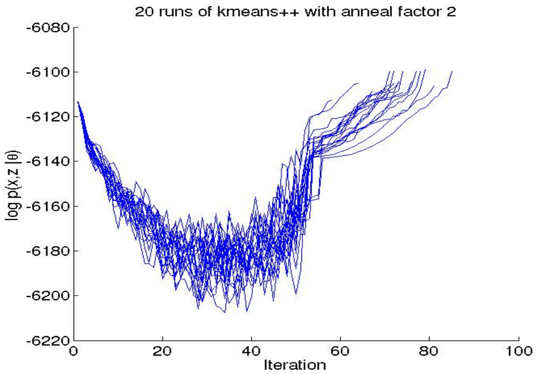

Figure 3: 20 runs

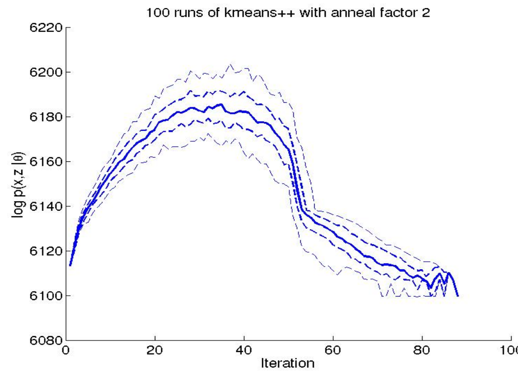

Figure 4: quantile for 100 runs

2.2) Approximation

For the initialization with Auxilary dishes, the **Aim** is to find the table,dish configuration for the first Restaurant to maximize the log likelihood:

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

$=$

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

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

W:number of unique words

$n^w_{..k}$number of occurence of word w in dish k

Below are the sub-problems:

1)$m_j = 1$:

merge it to Restaurant q

$\delta P = log(\frac{\Gamma(n_{j..}+n_{q..}+W\phi_0)}{\Gamma(n_{j..}+W\phi_0)\Gamma(n_{q..}+W\phi_0)}) + log(\Pi_{w=1}^{W}\frac{\Gamma(\phi_0+n^w_{j..})\Gamma(\phi_0+n^w_{q..})}{\Gamma(\phi_0+n^w_{j..}+n^w_{q..})})$

2)$m_j = 2$:

merge it to Restaurant p,q

$\delta P = \underline{log(m.. + \gamma) + log(\frac{\Gamma(n_{j2.}+\alpha)}{\Gamma(n_{j1.})\Gamma(n_{j2.})}) - log(\alpha)}$

$+$

$log(\frac{\Gamma(n_{j1.}+n_{p..}+W\phi_0)\Gamma(n_{j2.}+n_{q..}+W\phi_0)}{\Gamma(n_{j..}+W\phi_0)\Gamma(n_{p..}+W\phi_0)\Gamma(n_{q..}+W\phi_0)}) + log(\Pi_{w=1}^{W}\frac{\Gamma(\phi_0+n^w_{j..})\Gamma(\phi_0+n^w_{p..})\Gamma(\phi_0+n^w_{q..})}{\Gamma(\phi_0+n^w_{j1.}+n^w_{p..})\Gamma(\phi_0+n^w_{j2.}+n^w_{q..})})$

3)$m_j = 3$...

Strategy:

1. Exact Solution run time (that I can think of) is $\mathcal{O}(R^{m_j})\mathcal{O}(n^2_{j..})$
2. The bad global maxima mentioned above is led by t-term, since k-term is weak because of the uninformative auxilary dishes. we may simply drop the t-term.
3. Using Greedy approximation,the run time is $\mathcal{O}(R^{m_j})\mathcal{O}(n_{j..})$ or $\mathcal{O}(n_{j..})$

Figure 5: Initialization after pruning

Figure 6: Initialization with greedy approximation

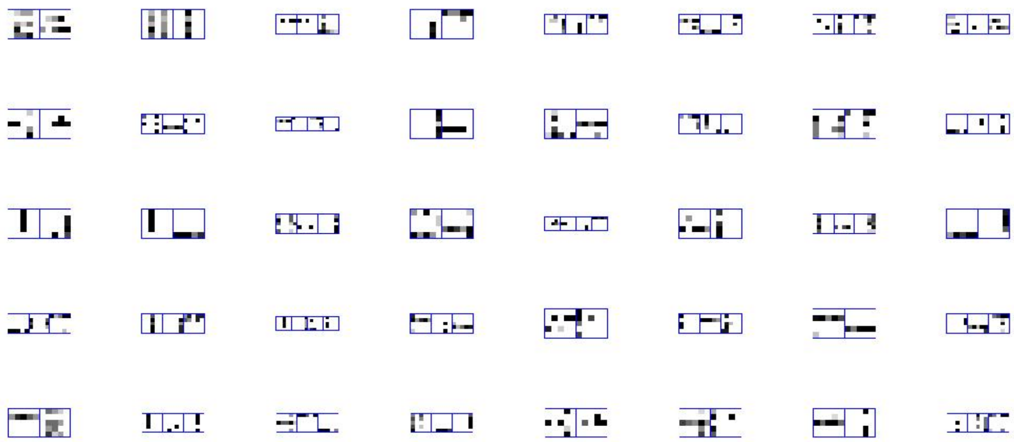

Figure 7: Split table after Initialization

3) Search

3.0) Statistics

In general, we are starting from initial guess of the statistics and want it move to the ground truth.

Sufficient Statistics: $\{t_{ji}\}, \{k_{jt}\}$ $\Longrightarrow$Local Search table/dish

Statistics: $\{m_{j\cdot}\}, K$ $\Longrightarrow$Split/Merge table/dish

3.1) Bigger Moves

Imitate the Initialization that:
1) Do the move statically instead of sequentially
2)

1. New config for tables in Restaurant j:Split all the tables in Restaurant j+local search t,k+merge t

2. New config for tables in Dish k: Split all the tables in Dish k+local search t,k

3. New config for dishes: Split all the dishes

3.2) Stochastic Search

Some split-table moves are bad simply because other similar tables in the same dish do not let it go. Instead of creating chaos by Bigger moves, simply we can give split move some probability to "tunnel" through the well.