14 September 2016

**Cluster Variation Method**

1. **Bethe-Peierls Approximation**

In this document, we will use *i*, *j*, … to label the flows as variable nodes, *a*, *b*, … to label the factor nodes at the vertices, and *α*, *β*, … to label the regions of factor nodes.

Denote the incidence matrix as for link *i* and vertex *a*. We adopt the following convention,



Hence defining *yi* to be the flow towards *a* = max{*b* | *i* ∈ *b*}, the flow conservation constraint becomes .

Note that 







Gradient with respect to :





Gradient with respect to :





To derive the cluster variation equations, we let

 and .

Compatibility between the probabilities at node *i* and cluster *a*:





Collecting those terms independent of  into an integral ,

 where



Considering the similar expression for the other vertex *b* of link *i*,



Multiplying,





Substituting into the expression for ,





1. **Generalized Belief Propagation**

**5**

**1**

***a***

***b***

***j***

***k***

**4**

**2**

**3**

**1**

***c***

***l***

**3**

***i***

**3**

***h***

**3**

***g***

**1**

**1**

***d***

***e***

**3**

***f***

**5**

**6**

**7**

***λ***

***μ***

***ν***

***ρ***

***σ***

***τ***

**8**

**9**

**10**

**11**

**12**

**13**

**14**

**15**

**16**

**17**

**18**

**19**

**20**

**21**

**22**

**23**

**24**

A network of square plaquettes. The circles represent the flows as the variable nodes located on edges, and the squares represent the factor nodes located at vertices. A cluster is centered at a factor node, and a region consists of 4 clusters each centered at a vertex of a square. Each variable node is a member of 2 clusters and 6 regions. For example, node 1 is a member of the clusters *k* and *l*, and is also a member of the regions *λ*, *μ*, *ν*, *ρ*, *σ* and *τ*. Each factor node is a member of 4 regions. For example, node *k* is a member of the regions *λ*, *μ*, *σ* and *τ*.

To calculate the total entropy in the cluster variation method for the network of square plaquettes, we first sum over the entropy of the regions. Since each factor node is a member of 4 regions, we have to subtract from it 3 times of the entropy of the clusters to give the correct counting of the factor nodes. Then in this entropy sum, since each variable node is a member of 2 clusters and 6 regions, it is counted 6 – (3)(2) = 0 times. Hence we have to add back the entropy of the nodes. In summary, the total entropy is the entropy of the regions, minus 3 times the entropy of the clusters, plus the entropy of the nodes.

Introduce the energy terms

, , 

Then the average energy is rewritten as





The free energy is given by

















Interchanging the order of summation in the red box,





Interchanging the order of summation in the blue box,





Interchanging the order of summation in the green box,





Summarizing, the free energy is given by

 where

























 where 

Gradient with respect to :





Gradient with respect to :





To derive the cluster variation equations, we let

,





***a***

***i***

***a***

***b***

***b***

***a***

***i***

***i***

***i***

***b***

***b***

***i***

***σ***

***σ***

***σ***

***σ***

**5**

***a***

***a***

***i***

***i***

***i***

**1**

***a***

***a***

***i***

***i***

***i***

***i***

***τ***

***σ***

***τ***

***τ***

***τ***

***τ***

***i***

***i***

***i***

***i***

***i***

***τ***

***τ***

***τ***

***b***

***b***

***b***

***b***

***b***

***b***

***b***

***b***

Schematic representations of the node, factor and region probabilities. Line arrows represent factor-to-node messages. Thick arrows represent region-to-factor messages.

Compatibility between the probabilities at node *i* and cluster *a*:









Next, we consider the compatibility between the probabilities of cluster *a* and region *σ*:









***a***

***i***

***b***

***a***

***j***

***j***

***i***

***b***

***b***

***j***

***σ***

***σ***

***σ***

***σ***

**∝**

***a***

***b***

***i***

***j***

***j***

**1**

***b***

***b***

***j***

***i***

***τ***

***σ***

***τ***

***τ***

***τ***

***τ***

***j***

***j***

***i***

***j***

***j***

***τ***

***τ***

***b***

***b***

***b***

***b***

***b***

***b***

***a***

***i***

***b***

***b***

***i***

***σ***

**∝**

Schematic representation of the generalized belief propagation equations.

1. **Generalized Belief Propagation in Triangular Lattices**

The cluster variation method for triangular lattices can be worked out similarly. In the triangular lattice, the variable nodes are located on the edges, and the factor nodes located at the vertices. A cluster is centered at a factor node, and a region consists of 3 clusters whose centers are connected by a triangular plaquette. Each variable node is a member of 2 clusters and 10 regions. Each factor node is a member of 6 regions.

To calculate the total entropy in the cluster variation method for the network of triangular plaquettes, we first sum over the entropy of the regions. Since each factor node is a member of 6 regions, we have to subtract from it 5 times of the entropy of the clusters to give the correct counting of the factor nodes. Then in this entropy sum, since each variable node is a member of 2 clusters and 10 regions, it is counted 10 – (5)(2) = 0 times. Hence we have to add back the entropy of the nodes. In summary, the total entropy is the entropy of the regions, minus 5 times the entropy of the clusters, plus the entropy of the nodes.















Gradient with respect to :



Gradient with respect to :







Gradient with respect to :





Let , ,

and .

Compatibility between the probabilities at node *i* and cluster *a*:







Collecting those terms independent of  into an integral ,

 where



For node *i* connected to clusters *a* and *b*, we have similarly



Multiplying,







This enables us to eliminate the sum of the cluster messages in the original equation, yielding



Similarly, for *b* = ∂*i*\*a*,



Hence in practice, the messages  and  can be calculated at the same time:





The flow distribution is given by



Next, we consider the compatibility between the probabilities of cluster *a* and region *σ*:









Note that for each *i* ∈ *σ*, either *i* ∈ *σ*\*a* or *i* ∈ *σ*∩*a*. Collecting those terms independent of  into an integral ,

 where



Similar expressions can be found for the other 5 regions to which *a* belongs. Multiplying,



This enables us to eliminate the sum of the region messages in the original equation, yielding



1. **Generalized Belief Propagation in Random Networks**

In random networks, the variable nodes are located on the edges, and the factor nodes located at the vertices. A cluster is centered at a factor node, and a region consists of clusters whose centers are connected by a plaquette. Each variable node is a member of 2 clusters. For a variable node *i* being a member of clusters *a* and *b*, we denote *ga* as the number of regions to which cluster *a* belongs, and *gi* as the number of regions to which both clusters *a* and *b* belong.

To calculate the total entropy in the cluster variation method for the random network, we first sum over the entropy of the regions. Since each factor node *a* is a member of *ga* regions, we have to subtract from it the sum of *ga* – 1 times of the entropy of cluster *a* to give the correct counting of the factor nodes. Then in this entropy sum, the counting of each variable node *i* is *ga* + *gb* – *gi* times in the entropy of the regions, and minus *ga* – 1 + *gb* – 1 times in the entropy of the clusters, and the total count is 2 – *gi*. Hence we have to add back *gi* – 1 times the entropy of node *i* to restore the counting of the variable nodes to 1. In summary,



















Gradients:













Let , ,

and .

Compatibility between node probabilities and cluster probabilities:







 where



Taking logarithm of both sides,



Since each node is a member of two clusters,









The flow distribution is given by



Check: For the square lattice, *ga* = *gb* = 4, *gi* = 2.

 OK

For a tree, *ga* = *gb* = *gi* = 0.

 OK

For a node connecting a tree and a loop, *gb* = *gi* = 0.





Compatibility between cluster probabilities and region probabilities:







 where



Taking logarithm of both sides,



Adding the same expressions for all {*σ* | *a* ∈ *σ*},



Substituting,

