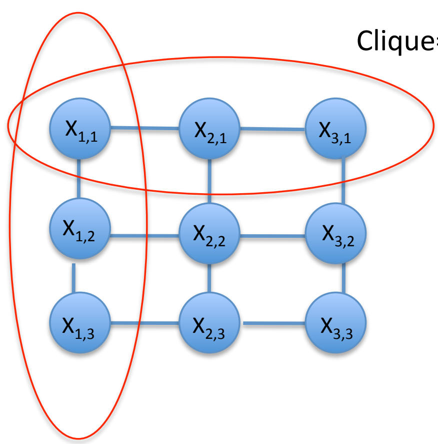
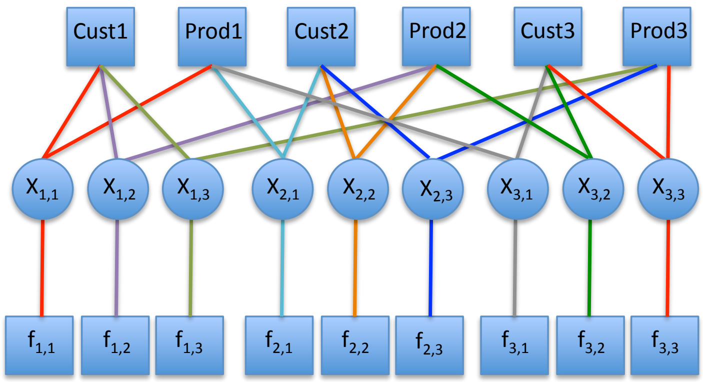
A Graphical Model Approach to the Assignment Problem

September 26, 2011

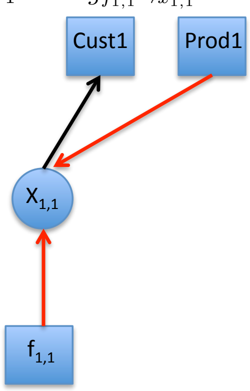
When the answer to the matching problem is unique, the problem can be formulated as a maximum a posteriori–MAP–problem in graphical models. In the graphical model, the variables x_{ij} are the assignment variables, ie indicator variables that indicate a match between a customer i and a product j, and SS_{ij} are the scores in matching customer i with product j. The joint distribution to optimize to is $p(\vec{x}) = exp(2 \cdot SS_{ij} \cdot x_{ij})\mathcal{I}_{\mathrm{cust}}\{(\mathrm{x_{ij}}\})\mathcal{I}_{\mathrm{prod}}\{(\mathrm{x_{ij}}\})$, which is similar to a Boltzmann distribution, but accounts for the constraint of only making one assignment per customer/product, by using the functions: $\mathcal{I}_{cust}\{(x_{ij}\}_{j=1}^{n}$, which equals 1 if $\sum_j x_{i,j}, = 1$, in other words, if a customer is only assigned to one product and equals 0 otherwise; $\mathcal{I}_{prod}\{(\mathrm{x_{ij}}\}_{\mathrm{i=1}}^{\mathrm{n}}) = 1$ if a product is assigned to only one customer and 0 otherwise. To formulate the graphical model, I first determine the cliques, which capture the relationship variables x_{ij} have with one another. I've identified two cliques for each customer: for example fixing $C = 1$: the set of variables $[x_{1,j}]\,\forall j = 1 \ldots n$ which indicates whether $cust_1$ is assigned to a product j, and the group of variables $[x_{i,1}]$, In the graphical model, the variables must be linked to reflect their clique membership.



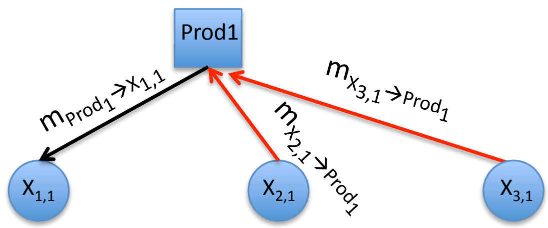
There is a potential function $\psi_C(x_C)$ associated with each clique. The joint distribution is the product of the clique potentials. To convert the undirected graph above into a factor graph, factorize the clique potentials into factors where each factor is a function of a corresponding set of variables in the clique. Since the joint distribution is $p(\vec{x}) = exp(SS_{ij} \cdot x_{ij})\mathcal{I}_{\mathrm{i}}\{(\mathrm{x_{ij}}\})\mathcal{I}_{\mathrm{j}}\{(\mathrm{x_{ij}}\})$, the two types of cliques can be factorized in the following way: for clique 1: $\psi_1(x_{1,1}, x_{1,2}x_{1,3} \ldots x_{1,n}) = \prod_{j=1}^{n} f(x_{1,j})\mathcal{I}_{cust}\{(\mathrm{x_{1,j}}\}_{\mathrm{j=1}}^{\mathrm{n}}))$ where $f(x_{1;j}) = exp(SS_{ij} \cdot x_{ij})$, $\psi_1(x_{1,1}, x_{2,1}, x_{3,1})$ can be factored as $\psi_1(x_{1,1}, x_{2,1}x_{3,1} \ldots x_{n,1}) = \prod_{i=1}^{n} f(x_{i,1})\mathcal{I}_{prod}\{(x_{i,1}\}_{i=1}^{n})$. The factor functions $f(x_{i,j})$, $I_{\mathrm{cust}}\{\mathrm{x_{ij}}\}$, and $\mathcal{I}_{prod}\{(\mathrm{x_{ij}}\}_{\mathrm{i=1}}^{\mathrm{n}}$ are represented as square factor nodes in the graphical model below (modelled on three customers and three products), and the variable nodes are the circle nodes.



In the graphical model above, the variable nodes are $x_{i,j}$, the factor nodes are $f(x_{i,j})$ and the indicator functions are represented by the factor nodes labelled "Customer" and "Product." To find the combination of variables $\vec{x}$ that maximize the joint distribution $p(\vec{x})$, I use the max product message passing algorithm. The message from a variable node to a factor node: ie $m_{x_{1,1}} \rightarrow cust_1$ is the product of incoming messages to $x_{1,1}$, that is, $m_{x_{1,1}} \rightarrow cust_1 = msg_{f_{1,1} \rightarrow x_{1,1}} \cdot msg_{prod_1 \rightarrow x_{1,1}}$.



The message from factor node Prod1 to variable $x_{1,1}$ is the function at the factor node times the product of incoming messages to Prod1: $msg_{Prod1 \rightarrow x_{1,1}} = msg_{x_{2,1} \rightarrow prod_1} \cdot msg_{x_{3,1} \rightarrow prod_1} \cdot I\{x_{i,1}\}_{i=1}^{n}$.



In my code, the vector of variables $\vec{x}$ converges when the decision rule returns the same vector 10 times in a row. The message algorithm is:

1. Initialize all messages to 1
2. $msg^{itr=k}_{x_{ij} \to cust_i}(x_{ij}) = f(x_{ij}) \cdot msg^{itr=k-1}_{prod_j \to x_{ij}}(x_{ij})$ where $x_{ij} = 0, 1$
3. $msg^{itr=k}_{x_{ij} \to prod_j}(x_{ij}) = f(x_{ij}) \cdot msg^{itr=k-1}_{cust_i \to x_{ij}}(x_{ij})$ where $x_{ij} = 0, 1$
4. $msg^{k}_{prod_j}(x_{ij}) = max_{\vec{x} \neq x_{ij}} \left\{ \mathcal{I}_{\mathrm{prod_j}} \{x_{ij}\}^n_{i=1} \prod_{m \neq i} msg^{k-1}_{x_{mj} \to prod_j}(x_{mj}) \right\}$
5. $msg^{k}_{cust_i}(x_{ij}) = max_{\vec{x} \neq x_{ij}} \left\{ \mathcal{I}_{\mathrm{cust_i}} \{x_{ij}\}^n_{j=1} \prod_{l \neq i} msg^{k-1}_{x_{il} \to cust_i}(x_{il}) \right\}$
6. Calculate belief $b^{itr=k}(x_{ij}) = f_{ij}(x_{ij}) \cdot msg^{k}_{cust_i \to x_{ij}}(x_{ij}) \cdot msg^{k}_{prod_j \to x_{ij}}(x_{ij})$ for $x^k_{ij} = 1$ and $x^k_{ij} = 0$
7. Decision rule: choose $x^k_{ij} = 1$ if $b(1) > b(0)$