Machine Learning 2 - Homework Week 3

Minh Ngo
MSc Artificial Intelligence
University of Amsterdam
nlminhtl@gmail.com

April 25, 2016

Collaborators: Arthur Bražinskas

Problem 2

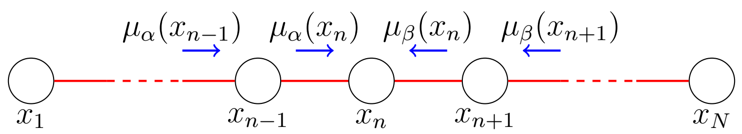

Figure 1: Chain of nodes model

To execute the sum-product algorithm the model described in the figure 1 is transformed to the factor graph 2.

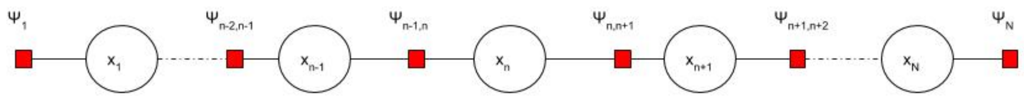

Figure 2: Factor graph

where $\psi_{n-1,n}$ is a factor from x_{n-1} to x_n variable. Applying Bishop (8.67, 8.66) we get:

$$\mu_{x_{n+1}\to\psi_{n,n+1}} = \mu_{\psi_{n+1,n+2}\to x_{n+1}}(x_{n+1}) \tag{1}$$

$$\begin{aligned}\underbrace{\mu_{\psi_{n,n+1}\to x_{n+1}}}_{\mu_\beta(x_n)} &= \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})\mu_{x_{n+1}\to\psi_{n,n+1}}(x_{n+1}) \\ &= \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})\underbrace{\mu_{\psi_{n+1,n+2}\to x_{n+1}}(x_{n+1})}_{\mu_\beta(x_{n+1})}\end{aligned} \tag{2}$$

$$\mu_{x_{n-1}\to\psi_{n-1,n}} = \mu_{\psi_{n-2,n-1}\to x_{n-1}}(x_{n-1}) \tag{3}$$

$$\begin{aligned}\underbrace{\mu_{\psi_{n-1,n}\to x_{n-1}}}_{\mu_\alpha(x_n)} &= \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n)\mu_{x_{n-1}\to\psi_{n-1,n}}(x_{n-1}) \\ &= \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n)\underbrace{\mu_{\psi_{n-2,n-1}\to x_{n-1}}(x_{n-1})}_{\mu_\alpha(x_{n-1})}\end{aligned} \tag{4}$$

Consequently:

$$\mu_\beta(x_n) = \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})\mu_\beta(x_{n+1}) \tag{5}$$

$$\mu_\alpha(x_n) = \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n)\mu_\alpha(x_{n-1}) \tag{6}$$

Setting Z to be a normalization constant.

$$p(x_n) = \frac{1}{Z}\mu_\alpha(x_n)\mu_\beta(x_n) \qquad q.e.d. \tag{7}$$

Problem 3

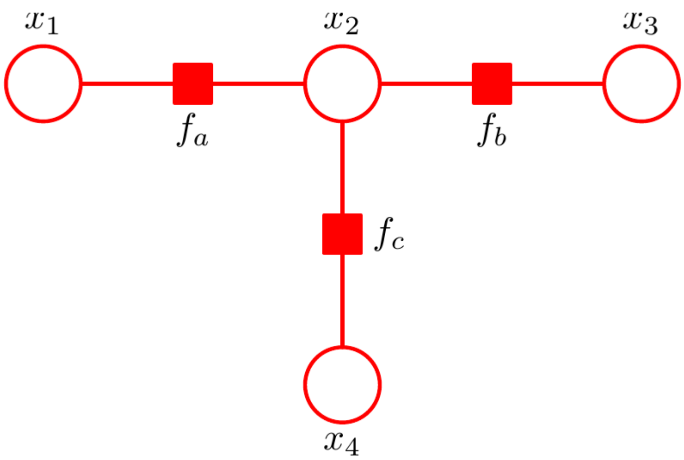

Figure 3: A simple factor graph

3.1

Applying again Bishop (8.66, 8.67):

$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3) \tag{8}$$

$$\mu_{f_c \to x_2}(x_2) = \sum_{x_4} f_c(x_2, x_4) \tag{9}$$

$$\mu_{x_2 \to f_a}(x_2) = \mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2) \tag{10}$$

$$\mu_{f_a \to x_1}(x_1) = \sum_{x_2} f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2) \tag{11}$$

$$\begin{aligned} p(x_1) &= \mu_{f_a \to x_1}(x_1) \\ &= \sum_{x_2} f_a(x_1, x_2) \sum_{x_3} f_b(x_2, x_3) \sum_{x_4} f_c(x_2, x_4) \\ &= \sum_{x_2} \sum_{x_3} \sum_{x_4} p(x_1, x_2, x_3, x_4) \end{aligned} \tag{12}$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2) \tag{13}$$

$$\mu_{x_2 \to f_b}(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_c \to x_2}(x_2) \tag{14}$$

$$\mu_{f_b \to x_3}(x_3) = \sum_{x_2} f_b(x_2, x_3)\mu_{x_2 \to f_b}(x_2) \tag{15}$$

$$\begin{aligned} p(x_3) &= \mu_{f_b \to x_3}(x_3) \\ &= \sum_{x_2} f_b(x_2, x_3) \sum_{x_1} f_a(x_1, x_2) \sum_{x_4} f_c(x_2, x_4) \\ &= \sum_{x_2} \sum_{x_1} \sum_{x_4} p(x_1, x_2, x_3, x_4) \end{aligned} \tag{16}$$

3.2

Applying Bishop (8.72) we get:

$$\begin{aligned} p(x_1, x_2) &= f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2)\mu_{x_1 \to f_a}(x_1) \\ &= f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2) \\ &= f_a(x_1, x_2) \sum_{x_3} f_b(x_1, x_2) \sum_{x_4} f_c(x_2, x_4) \\ &= \sum_{x_3} \sum_{x_4} p(x_1, x_2, x_3, x_4) \end{aligned} \tag{17}$$