CMPSCI 688 : Graphical Models
Assignment 1

Jeevan Shankar
jshankar@cs.umass.edu

February 10, 2014

1. Factorization

$$P(X) = \prod_{i=1}^{\#nodes} P(X_i | Pa_{X_i}^G)$$

$$P(A, G, CH, BP, HD, HR, CP, EIA, ECG) = P(A)P(G)(CH|A, G)P(BP|G)P(HD|CH, BP)$$
$$\times P(HR|A, BP, HD)P(CP|HD)P(EIA|HD)P(ECG|HD)$$

2. Likelihood Function

$$\begin{aligned}
L(\theta) &= \frac{1}{N}\sum_{n=1}^{N} \log P(X = x_n) \\
&= \frac{1}{N}\sum_{n=1}^{N} \log P(A = a_n, G = g_n, CH = ch_n, BP = bp_n, HD = hd_n, HR = hr_n, CP = cp_n, \\
&\quad EIA = eia_n, ECG = ecg_n) \\
&= \frac{1}{N}\sum_{n=1}^{N} \log P(A = a_n) + \log P(G = g_n) + \log P(CH = ch_n|A = a_n, G = g_n) + \log P(BP = bp_n|G = g_n) \\
&\quad + \log P(HD = hd_n|CH = ch_n, BP = bp_n) + \log P(HR = hr_n|A = a_n, BP = bp_n, HD = hd_n) + \\
&\quad + \log P(CP = cp_n|HD = hd_n) + \log P(EIA = eia_n|HD = hd_n) + \log P(ECG = ecg_n|HD = hd_n) \\
&= \frac{1}{N}\sum_{n=1}^{N} \sum_{a} [a = a_n] \log P(A = a) + \sum_{g} [g = g_n] \log P(G = g) \\
&\quad + \sum_{ch,a,g} [ch = ch_n, a = a_n, g = g_n] \log P(CH = ch|A = a, G = g) + \sum_{bp,g} [bp = bp_n, g = g_n] \log P(BP = bp|G = g) \\
&\quad + \sum_{hd,ch,bp} [hd = hd_n, ch = ch_n, bp = bp_n] \log P(HD = hd|CH = ch, BP = bp) \\
&\quad + \sum_{hr,a,bp,hd} [hr = hr_n, a = a_n, bp = bp_n, hd = hd_n] \log P(HR = hr|A = a, BP = bp, HD = hd) \\
&\quad + \sum_{eia,hd} [eia = eia_n, hd = hd_n] \log P(EIA = eia|HD = hd) + \sum_{ecg,hd} [ecg = ecg_n, hd = hd_n] \log P(ECG = ecg|HD = hd) \\
&\quad + \sum_{cp,hd} [cp = cp_n, hd = hd_n] \log P(CP = cp|HD = hd)
\end{aligned}$$

$$\begin{aligned}
L(\theta) &= \frac{1}{N}\sum_{n=1}^{N} \sum_{a} [a = a_n] \log \theta_a^A + \sum_{g} [g = g_n] \log \theta_g^G + \sum_{ch,a,g} [ch = ch_n, a = a_n, g = g_n] \log \theta_{ch|a,g}^{CH} + \sum_{bp,g} [bp = bp_n, g = g_n] \log \theta_{bp|g}^{BP} \\
&\quad + \sum_{hd,ch,bp} [hd = hd_n, ch = ch_n, bp = bp_n] \log \theta_{hd|ch,bp}^{HD} + \sum_{hr,a,bp,hd} [hr = hr_n, a = a_n, bp = bp_n, hd = hd_n] \log \theta_{hr|a,bp,hd}^{HR} \\
&\quad + \sum_{cp,hd} [cp = cp_n, hd = hd_n] \log \theta_{cp|hd}^{CP} + \sum_{eia,hd} [eia = eia_n, hd = hd_n] \log \theta_{eia|hd}^{EIA} + \sum_{ecg,hd} [ecg = ecg_n, hd = hd_n] \log \theta_{ecg|hd}^{ECG}
\end{aligned}$$

3. Maximum Likelihood Estimators

Forming the Lagrangian, we get

$$
\begin{aligned}
L(\theta, \lambda) = \frac{1}{N} \sum_{n=1}^{N} & \left(\sum_{a} [a = a_n] \log \theta_a^A \right) - \lambda^A \left(\sum_{a} \theta_a^A - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{g} [g = g_n] \log \theta_g^G \right) - \lambda^G \left(\sum_{g} \theta_g^G - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{ch,a,g} [ch = ch_n, a = a_n, g = g_n) \log \theta_{ch|a,g}^{CH} \right) - \sum_{a,g} \lambda_{a,g}^{CH} \left(\sum_{ch} \theta_{ch|a,g}^{CH} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{bp,g} [bp = bp_n, g = g_n] \log \theta_{bp|g}^{BP} \right) - \sum_{g} \lambda_g^{BP} \left(\sum_{bp} \theta_{bp|g}^{BP} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{hd,ch,bp} [hd = hd_n, ch = ch_n, bp = bp_n] \log \theta_{hd|ch,bp}^{HD} \right) - \sum_{ch,bp} \lambda_{ch,bp}^{HD} \left(\sum_{hd} \theta_{hd|bp,ch}^{HD} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{hr,a,bp,hd} [hr = hr_n, a = a_n, bp = bp_n, hd = hd_n] \log \theta_{hr|a,bp,hd}^{HR} \right) - \sum_{a,bp,hd} \lambda_{a,bp,hd}^{HR} \left(\sum_{hr} \theta_{hr|a,bp,hd}^{HR} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{cp,hd} [cp = cp_n, hd = hd_n] \log \theta_{cp|hd}^{CP} \right) - \sum_{hd} \lambda_{hd}^{CP} \left(\sum_{cp} \theta_{cp|hd}^{CP} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{eia,hd} [eia = eia_n, hd = hd_n] \log \theta_{eia|hd}^{CP} \right) - \sum_{hd} \lambda_{hd}^{EIA} \left(\sum_{eia} \theta_{eia|hd}^{EIA} - 1 \right) \\
& + \frac{1}{N} \sum_{n=1}^{N} \left(\sum_{ecg,hd} [ecg = ecg_n, hd = hd_n] \log \theta_{ecg|hd}^{CP} \right) - \sum_{hd} \lambda_{hd}^{ECG} \left(\sum_{ecg} \theta_{ecg|hd}^{ECG} - 1 \right)
\end{aligned}
$$

Now, to get MLE estimate of $\theta^{HR}_{L|1,H,Y}$, we put $\dfrac{\partial L(\theta,\lambda)}{\partial \theta^{HR}_{L|1,H,Y}} = 0$ and $\dfrac{\partial L(\theta,\lambda)}{\partial \theta^{HR}_{H|1,H,Y}} = 0$

$$\frac{1}{N} \sum_{n=1}^{N} [hr_n = L, a_n = 1, bp_n = H, = Y] \frac{1}{\theta^{HR}_{L|1,H,Y}} - \lambda^{HR}_{1,H,Y} = 0 \tag{1}$$

$$\frac{1}{N} \sum_{n=1}^{N} [hr_n = H, a_n = 1, bp_n = H, = Y] \frac{1}{\theta^{HR}_{H|1,H,Y}} - \lambda^{HR}_{1,H,Y} = 0 \tag{2}$$

To eliminate $\lambda^{HR}_{H|1,H,Y}$, we can use the constraint $\theta^{HR}_{L|1,H,Y} + \theta^{HR}_{H|1,H,Y} = 1$

$$\frac{1}{N} \sum_{n=1}^{N} [hr_n = L, a_n = 1, bp_n = H, hd_n = Y] \frac{1}{\lambda^{HR}_{1,H,Y}} + \frac{1}{N} \sum_{n=1}^{N} [hr_n = H, a_n = 1, bp_n = H, hd_n = Y] \frac{1}{\lambda^{HR}_{1,H,Y}} = 1$$

$$\frac{1}{N} \sum_{n=1}^{N} [a_n = 1, bp_n = H, hd_n = Y] \frac{1}{\lambda^{HR}_{1,H,Y}} = 1 \tag{3}$$

Now, substitution (3) in (1), we get

$$
\begin{aligned}
\frac{1}{N} \sum_{n=1}^{N} [hr_n = L, a_n = 1, bp_n = H, = Y] \frac{1}{\theta^{HR}_{L|1,H,Y}} - \frac{1}{N} \sum_{n=1}^{N} [a_n = 1, bp_n = H, hd_n = Y] = 0 \\
\theta^{HR}_{L|1,H,Y} = \frac{\sum_{n=1}^{N} [hr_n = L, a_n = 1, bp_n = H, hd_n = Y]}{\sum_{n=1}^{N} [a_n = 1, bp_n = H, hd_n = Y]}
\end{aligned}
$$

Thus,

$$\boxed{\theta^{HR}_{L|1,H,Y} = \frac{\sum_{n=1}^{N} [hr_n = L, a_n = 1, bp_n = H, hd_n = Y]}{\sum_{n=1}^{N} [a_n = 1, bp_n = H, hd_n = Y]}}$$

4. Learning

P(A)	A
0.185	< 45
0.268	45-55
0.547	≥ 55

P(BP|G)	BP	G
0.333	Low	Female
0.667	High	Female
0.476	Low	Male
0.524	High	Male

P(HD|BP, CH)	HD	BP	CH
0.5	N	Low	Low
0.5	Y	Low	Low
0.556	N	High	Low
0.444	Y	High	Low
0.576	N	Low	High
0.424	Y	Low	High
0.492	N	High	High
0.508	Y	High	High

P(HR|A, BP, HD)	HR	A	BP	HD
0.056	Low	< 45	Low	N
0.944	High	< 45	Low	N
0.227	Low	$45 - 55$	Low	N
0.773	High	$45 - 55$	Low	N
0.526	Low	≥ 55	Low	N
0.474	High	≥ 45	Low	N
0.071	Low	< 45	High	N
0.929	High	< 45	High	N
0.174	Low	$45 - 55$	High	N
0.826	High	$45 - 55$	High	N
0.219	Low	≥ 55	High	N
0.781	High	≥ 45	High	N
0.5	Low	< 45	Low	Y
0.5	High	< 45	Low	Y
0.462	Low	$45 - 55$	Low	Y
0.538	High	$45 - 55$	Low	Y
0.609	Low	≥ 55	Low	Y
0.391	High	≥ 45	Low	Y
0.667	Low	< 45	High	Y
0.333	High	< 45	High	Y
0.429	Low	$45 - 55$	High	Y
0.571	High	$45 - 55$	High	Y
0.525	Low	≥ 55	High	Y
0.475	High	≥ 45	High	Y

5. Probability Queries

Part (a)

$$P(CH = L|A = 2, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no) =$$

$$\frac{P(CH = Low, A = 2, G = M, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no)}{P(A = 2, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no)} =$$

$$\frac{P(CH = L, A = 2, G = M, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no)}{\sum_{ch\in(L,H)} P(CH = ch, A = 2, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no} =$$

$$\frac{P(CH = L|A = 2, G = M)P(HD = L|CH = L, BP = L)}{\sum_{ch\in(L,H)} P(CH = ch|A = 2, G = M)P(HD = L|CH = ch, BP = L)} =$$

(After applying factorization to Joint Probability)

$$0.135$$

(Looking up in CPTs of CH, HD)

Therefore, $P(CH = L|A = 2, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no)) = 0.135$ and $P(CH = H|A = 2, G = M, CP = None, BP = L, ECG = normal, HR = L, EIA = no, HD = no)) = 0.865$

Part (b)

$$P(BP = L|A = 2, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no) =$$

$$\frac{P(BP = L, A = 2, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no)}{\sum_{bp} P(BP = bp, A = 2, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no)} =$$

$$\frac{\sum_{g} P(BP = L, A = 2, G = g, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no)}{\sum_{bp}\sum_{g} P(BP = bp, A = 2, G = g, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no)} =$$

$$\frac{\sum_{g} P(G = g)P(CH = H|G = g, A = 2)P(BP = L|G = g)P(HR = H|A = 2, BP = L, HD = no)P(HD = no|BP = L, CH = H)}{\sum_{bp}\sum_{g} P(G = g)P(CH = H|G = g, A = 2)P(BP = bp|G = g)P(HR = H|A = 2, BP = bp, HD = no)P(HD = no|BP = bp, CH = H)}$$

(After applying factorization to Joint Probability and Simplifying)

$$= 0.455$$

(Looking up in the CPTs)

Therefore, $P(BP = L|A = 2, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no) = 0.455$ and $P(BP = H|A = 2, CP = Typical, CH = H, ECG = normal, HR = H, EIA = yes, HD = no) = 0.545$

6. Classification

Part (a) :

Part (b) :

$P(HD = N|A = a, G = g, CH = ch, BP = bp, HR = hr, CP = cp, EIA = eia, ECG = ecg) =$

$$\frac{\begin{array}{c}P(A{=}a)P(G{=}g)P(CH{=}ch|A{=}a,G{=}g)P(BP{=}bp|G{=}g)P(HD{=}N|CH{=}ch,BP{=}bp)\times\\ P(HR{=}hr|A{=}a,BP{=}bp,HD{=}N)P(CP{=}cp|HD{=}N)P(EIA{=}eia|HD{=}N)P(ECG{=}ecg|HD{=}N)\end{array}}{\sum_{hd}\begin{array}{c}P(A{=}a)P(G{=}g)P(CH{=}ch|A{=}a,G{=}g)P(BP{=}bp|G{=}g)P(HD{=}hd|CH{=}ch,BP{=}bp)\times\\ P(HR{=}hr|A{=}a,BP{=}bp,HD{=}hd)P(CP{=}cp|HD{=}hd)P(EIA{=}eia|HD{=}hd)P(ECG{=}ecg|HD{=}hd)\end{array}}$$

$$\frac{\begin{array}{c}\cancel{P(A{=}a)}\,\cancel{P(G{=}g)}\,\cancel{P(CH{=}ch|A{=}a,G{=}g)}\,\cancel{P(BP{=}bp|G{=}g)}\,P(HD{=}N|CH{=}ch,BP{=}bp)\times\\ P(HR{=}hr|A{=}a,BP{=}bp,HD{=}N)P(CP{=}cp|HD{=}N)P(EIA{=}eia|HD{=}N)P(ECG{=}ecg|HD{=}N)\end{array}}{\sum_{hd}\begin{array}{c}\cancel{P(A{=}a)}\,\cancel{P(G{=}g)}\,\cancel{P(CH{=}ch|A{=}a,G{=}g)}\,\cancel{P(BP{=}bp|G{=}g)}\,P(HD{=}hd|CH{=}ch,BP{=}bp)\times\\ P(HR{=}hr|A{=}a,BP{=}bp,HD{=}hd)P(CP{=}cp|HD{=}hd)P(EIA{=}eia|HD{=}hd)P(ECG{=}ecg|HD{=}hd)\end{array}}$$

$$\frac{P(HD{=}N|CH{=}ch,BP{=}bp)P(HR{=}hr|A{=}a,BP{=}bp,HD{=}N)P(CP{=}cp|HD{=}N)P(EIA{=}eia|HD{=}N)P(ECG{=}ecg|HD{=}N)}{\sum_{hd} P(HD{=}hd|CH{=}ch,BP{=}bp)P(HR{=}hr|A{=}a,BP{=}bp,HD{=}hd)P(CP{=}cp|HD{=}hd)P(EIA{=}eia|HD{=}hd)P(ECG{=}ecg|HD{=}hd)}$$

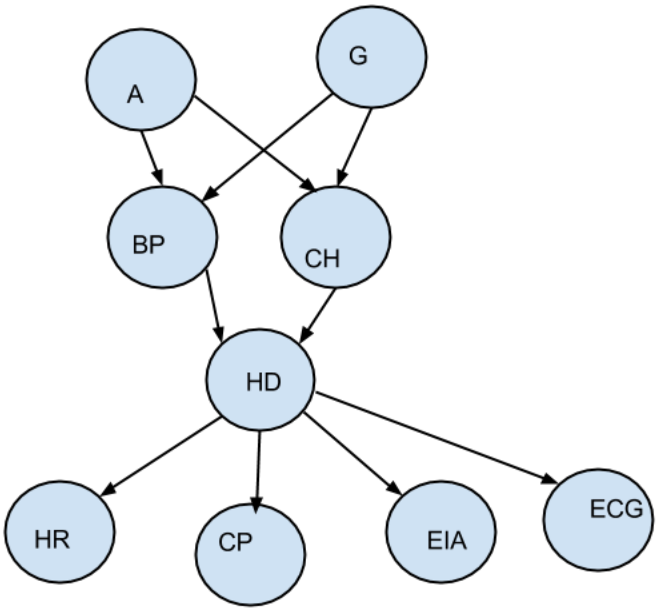
Part (c) :

Fold	No correct	No total	$\%Accuracy$
1	48	60	80
2	48	60	80
3	43	60	71.66
4	43	60	71.66
5	44	60	73.33
Avg(Mean)	45.2	60	75.33

The variance is 3.862.

7. Modelling

Part(a) : model



Part (b) : factorization

$$P(X) = \prod_{i=1}^{\#nodes} P(X_i|Pa_{X_i}^G)$$

$$P(A, G, CH, BP, HD, HR, CP, EIA, ECG) = P(A)P(G)(CH|A, G)P(BP|A, G)P(HD|CH, BP)$$
$$\times P(HR|HD)P(CP|HD)P(EIA|HD)P(ECG|HD)$$

Part (c) : design choice

Idea was to kind of simulate neural-network style like structure with no latent variables.

Part (d) : results

Fold	No correct	No total	$\%Accuracy$
1	48	60	80
2	47	60	78.33
3	45	60	75.00
4	43	60	71.66
5	42	60	70
Avg(Mean)	45	60	75

The Std Dev is 3.801.

Analysis: The average accuracy and Std Dev is slightly less than given model. Above model fairs better only in 3rd test data.