统计学习：第十一章

1 写出图 11.3 中无向图描述的概率图模型的因子分解式。
答:

$$P(Y_1, Y_2, Y_3, Y_4) = \frac{1}{Z}\Psi_{123}(Y_1, Y_2, Y_3)\Psi_{234}(Y_2 Y_3 Y_4)$$

2 证明 $Z(x) = \alpha_n^T(x)1 = 1^T\beta_1(x)$，其中 1 是元素均为 1 的 m 维列向量。
答:

$$\begin{aligned}
&\alpha_n^T(x)1 \\
=&\alpha_0^T(x)M_1(x)\dots M_n(x)1 \\
=&1^T M_1(x)\dots M_n(x)1 \\
=&\sum_y exp(w \cdot F(y,x)) \\
=&Z(x)
\end{aligned}$$

$$\begin{aligned}
&1^T\beta_1(x) \\
=&1^T M_1(x)\dots M_n(x)1 \\
=&Z(x)
\end{aligned}$$

3 写出条件随机场模型学习的梯度下降法。
答:

$$L = \sum_{j=1}^{N}(\sum_{k=1}^{K} w_k f_k(y_j, x_j) - Z_w(x_j))$$

为使损失函数最小，寻找梯度

$$\frac{\partial L}{\partial w_k} = \sum_{j=1}^{N}(f_k(y_j, x_j) - \sum_y \exp(w_k f_k(y_j, x_j))f_k(y_j, x_j))$$

故若采用随机梯度下降法, 学习率为 η, 更新为

$$w_{k,new} = w_{k,old} - \eta((f_k(y_j, x_j) - \sum_y \exp(w_k f_k(y_j, x_j))f_k(y_j, x_j)))$$

4 参考图 11.6 的状态路径图，假设随机矩阵 $M_1(x), M_2(x), M_3(x), M_4(x)$ 分别是

$$M_1 = \begin{bmatrix} 0 & 0 \\ 0.5 & 0.5 \end{bmatrix}, M_2 = \begin{bmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{bmatrix},$$

$$M_3 = \begin{bmatrix} 0.5 & 0.5 \\ 0.6 & 0.4 \end{bmatrix}, M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

求以 start = 2 为起点，以 stop = 2 为终点的所有路径的状态序列 y 的概率及概率最大的状态序列

答：非规范化概率为：

$y_{21112} = 0.5 \times 0.3 \times 0.5 = 0.075$

$y_{21122} = 0.5 \times 0.3 \times 0.5 = 0.075$

$y_{21212} = 0.5 \times 0.7 \times 0.6 = 0.21$

$y_{21222} = 0.5 \times 0.7 \times 0.4 = 0.14$

$y_{22112} = 0.5 \times 0.7 \times 0.5 = 0.175$

$y_{22122} = 0.5 \times 0.7 \times 0.5 = 0.175$

$y_{22212} = 0.5 \times 0.3 \times 0.6 = 0.09$

$y_{22222} = 0.5 \times 0.3 \times 0.4 = 0.06$

概率最大的为 $2, 1, 2, 1, 2$。