Sheet 6

Alexander Bigalke, Arthur Heimbrecht, Robin Rombach

July 4, 2018

⇒ Solutions starting next page...

Exercise 1

Exercise 1

Let's first make up some arbitrarily chosen probabilities. We chose:

$$\left.\begin{array}{l} p(C=0) = 0{,}7 \quad , \quad p(C=1) = 0{,}3 \\ p(A=0|C=0) = 0{,}4 \, , \quad p(A=1|C=0) = 0{,}6 \\ p(B=0|C=0) = 0{,}7 \, , \quad p(B=1|C=0) = 0{,}3 \\ p(A=0|C=1) = 0{,}5 \, , \quad p(A=1|C=1) = 0{,}5 \\ p(B=0|C=1) = 0{,}2 \, , \quad p(B=1|C=1) = 0{,}8 \end{array}\right\} \text{each line sums up to 1 as required by normalization}$$

Based on these probabilities, we can now create 2 tables for the conditional probability $p(A,B|C)$ for $C=0$ and $C=1$. Assuming $A \perp B \mid C$, we can use

$$p(A,B|C) = p(A|C) \cdot p(B|C)$$

to obtain

$C=0$

A	B	p(A|C)	p(B|C)	p(A,B|C)
0	0	0,4	0,7	0,28
0	1	0,4	0,3	0,12
1	0	0,6	0,7	0,42
1	1	0,6	0,3	0,18

$C=1$

A	B	p(A|C)	p(B|C)	p(A,B|C)
0	0	0,5	0,2	0,1
0	1	0,5	0,8	0,4
1	0	0,5	0,2	0,1
1	1	0,5	0,8	0,4

Now, we can marginalize over C to compute

$$p(A) = p(A|C=0) \cdot p(C=0) + p(A|C=1) \cdot p(C=1) \qquad \text{(analogously for } p(B)\text{)}$$

and

$$p(A,B) = p(A,B|C=0) \cdot p(C=0) + p(A,B|C=1) \cdot p(C=1)$$

$\Rightarrow$

A	B	p(A)	p(B)	p(A,B)		p(A) · p(B)
0	0	0,43	0,55	0,226	$\neq$	0,2365
0	1	0,43	0,45	0,204	$\neq$	0,1935
1	0	0,57	0,55	0,324	$\neq$	0,3135
1	1	0,57	0,45	0,246	$\neq$	0,2565

As one can see, $p(A,B) \neq p(A) \cdot p(B)$ in all 4 cases and thus A and B are not independent $(A \not\perp B)$ although $A \perp B \mid C$ holds

$\Rightarrow A \perp B \mid C$ does not imply $A \perp C$!

Exercise 2

2.1

2.1)

1. no additional information

$p(A=1) = p(A=\text{boy}) = \frac{1}{2}$ same for B

$\Rightarrow p(A=1, B=1) = p(A=1) \cdot p(B=1) = \frac{1}{4}$

2. we know one child is a boy

$\Rightarrow$ new conditional prob.

$$p(A=1 \wedge B=1 \mid A=1 \vee B=1) = \frac{p(A=1, B=1)}{p(A=1 \vee B=1)}$$

w/ $p(A \vee B) = p(\neg(\bar{A} \wedge \bar{B})) = 1 - p(\bar{A} \wedge \bar{B})$

$$= \frac{p(A=1)\, p(B=1)}{1 - p(A \neq 1\ B \neq 1)} = \frac{p(A=1)\, p(B=1)}{1 - p(A \neq 1)\, p(B \neq 1)}$$

$$= \frac{p(A=1)\, p(B=1)}{1 - (1 - p(A=1))(1 - p(B=1))} = \frac{1/2 \cdot 1/2}{1 - \frac{1}{2} \cdot 1/2} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

3. $p(A=1 \wedge B=1 \mid A=1) = \frac{p(A=1)\, p(B=1)}{p(A=1)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$

4. use $\frac{2C-1}{4C-1} = p$ w/ $C = 7$

$\Rightarrow p\,(A=1 \wedge B=1 \wedge (A_{Day} = \text{Sun} \vee B_{Day} = \text{Sun}) \mid (A=1 \wedge A_{Day} = \text{Sun}) \vee (B=1 \wedge B_{Day} =$

$= \frac{14-1}{28-1} = \frac{13}{27}$

5. $p = \frac{729}{1459}$ w/ $C = 365$

proof: assume we observe the arbitrary C^{th} possible state of

C no. of states:

$\Rightarrow p(A_c = C) = \frac{1}{C}$ same for B

$$\Rightarrow p = p\big(A=1 \wedge B=1 \wedge (A_c = C \vee B_c = C) \,\big|\, (A=1 \wedge A_c = C) \vee (B=1 \wedge B_c = C)\big)$$

$$= \frac{p(A=1 \wedge B=1 \wedge (A_c = C \vee B_c = C))}{p((A=1 \wedge A_c = C) \vee (B=1 \wedge B_c = C))} = \frac{p(A=1)\, p(B=1)\, p(A_c = C \vee B_c = C)}{1 - p(\neg(A=1 \wedge A_c = C) \wedge \neg(B=1 \wedge B_c = C))}$$

$$= \frac{p(A=1)\, p(B=1)\, (1 - p(A_c \neq C \wedge B_c \neq C))}{1 - (1 - p(A=1 \wedge A_c = C))(1 - p(B=1 \wedge B_c = C))}$$

$$= \frac{p(A=1)\, p(B=1)\, (1 - (1 - p(A_c = C))(1 - p(B_c = C)))}{1 - (1 - p(A=1)\, p(A_c = C))(1 - p(B=1)\, p(B_c = C))}$$

$$= \frac{1/2 \cdot 1/2 \cdot (1 - (1 - \frac{1}{C})(1 - \frac{1}{C}))}{1 - (1 - 1/2 \cdot \frac{1}{C})(1 - 1/2 \cdot 1/C)}$$

$$= \frac{\frac{1}{4} \cdot (1 - 1 + \frac{2}{C} - \frac{1}{C^2})}{1 - 1 + \frac{1}{C} - \frac{1}{4}\frac{1}{C^2}} = \frac{\frac{1}{4}\frac{2C-1}{C^2}}{\frac{4C-1}{4C^2}} = \frac{2C-1}{4C-1}$$

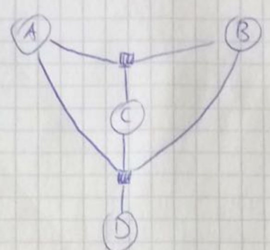
2.2

$\Rightarrow$ see the attached file `boys.html` or `boys.ipynb`.

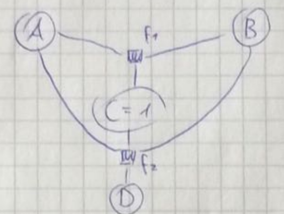
2.3

2.3)

factor graph:



factor graph $C=1$:



with $C=1$ the graph is no longer a loop but a tree such that belief propagation becomes possible

A	B	C	D	f_1	f_2
1	1	1	1	$2/3$	$1/4$
1	0	1	0	$2/3$	$3/4$
0	1	1	0	$2/3$	$3/4$
0	0	0	0	$1/3$	$3/4$

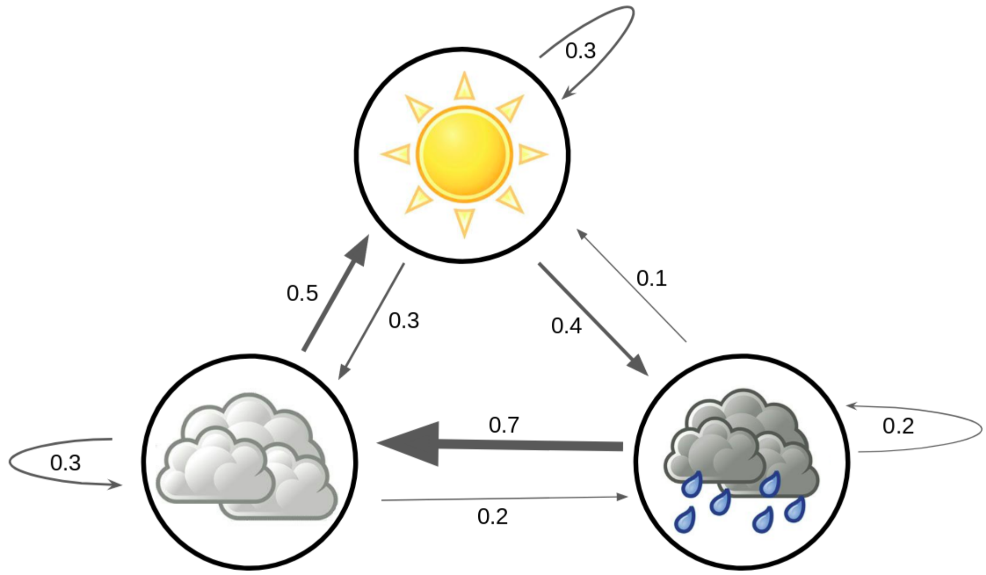
other comb. not possible

$\Rightarrow$ Round 1: $\mu_{C\to f_1} = 1 \qquad \mu_{C\to f_2} = 1$

Round 2: $\mu_{f_1\to A} = \frac{2}{3} \cdot 1 = \frac{2}{3} \quad \Rightarrow p(A \mid C=1) = 2/3$

Round 3: $\mu_{A\to f_2} = \frac{2}{3} \qquad p(D \mid C=1) = \frac{1}{3}$

Round 4: $\mu_{f_2\to D} = 3 \cdot f_2 \cdot (\underbrace{\mu_{C\to f_2}}_{=1} \underbrace{\mu_{B\to f_2}}_{=2/3} \underbrace{\mu_{A\to f_2}}_{=2/3}) = 3 \cdot \frac{1}{4} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{1}{3}$



The graph required for exercise *b)*

Exercise 3 - Weather Forecast using a Markov Chain

```
In [1]: import numpy as np
```

The task is to classify the weather on an arbitrary day n, where the weather is specified by $t_n \in \{rainy, cloudy, sunny\}$.

a) The transition probabilities from day $n-1$ to day n are given as follows:

$$p_{r,r} = 0.2\,;\ p_{c,r} = 0.7\,;\ p_{s,r} = 0.1$$
$$p_{r,c} = 0.2\,;\ p_{c,c} = 0.3\,;\ p_{s,c} = 0.5$$
$$p_{r,s} = 0.4\,;\ p_{c,s} = 0.3\,;\ p_{s,s} = 0.3$$

where we introduced the shorthand-notation $p_{i,j} := p\left(t_n = i \,\middle|\, t_{n-1} = j\right)$ and the abbreviations $r = rainy$, $s = sunny$ and $c = cloudy$.

The missing conditional probabilities were computed by using the condition:

$$\sum_i p_{i,j} = 1$$

b) see Figure 1.

c) Consider a given probability vector

$$\mathbf{p}(t_0) = \begin{pmatrix} p(t_0 = r) \\ p(t_0 = c) \\ p(t_0 = s) \end{pmatrix}$$

The probabilities after one day, i.e. $\mathbf{p}(t_0)$ are then given by

$$p(t_1 = j) = \sum_i p_{j,i} \cdot p(t_0 = i)$$

d) The above formula can be cast into matrix form:

$$\mathbf{p}(t_1) = \mathcal{P}^{\mathrm{T}}\,\mathbf{p}(t_0) \equiv \mathbf{M}\,\mathbf{p}(t_0)$$

by defining the transition matrix

$$\mathcal{P} = \begin{pmatrix} p_{r,r} & p_{c,r} & p_{s,r} \\ p_{r,c} & p_{c,c} & p_{s,c} \\ p_{r,s} & p_{c,s} & p_{s,s} \end{pmatrix} = \begin{pmatrix} 0.2 & 0.7 & 0.1 \\ 0.2 & 0.3 & 0.5 \\ 0.4 & 0.3 & 0.3 \end{pmatrix}$$

The numerical computation for $\mathbf{p}(t_0) = (0.5, 0.25, 0.25)^{\mathrm{T}}$ is demonstrated below.

```
In [2]: # numerical computation:
        P = np.array([[0.2, 0.7, 0.1], [0.2, 0.3, 0.5], [0.4, 0.3, 0.3]])
        p0 = np.array([0.5, 0.25, 0.25])
        M = P.T
        print('p1: ', M@p0)
```

```
p1:  [ 0.25  0.5   0.25]
```

Hence, $\mathbf{p}(t_1) = \mathcal{P}^{\mathrm{T}}\,\mathbf{p}(t_0) = (0.25, 0.5, 0.25)^{\mathrm{T}}$.

e) The above formula can be easily generalized to arbitrary times:

$$\mathbf{p}(t_n) = \mathbf{M}^n\,\mathbf{p}(t_0)$$

with $n \geq 1$.

```
In [5]: print('p100: ', np.linalg.matrix_power(P.T, 100) @ p0)
```

```
p100:  [ 0.265625  0.40625   0.328125]
```

Hence, $\mathbf{p}(t_{100}) = \mathcal{P}^{\mathrm{T}}\,\mathbf{p}(t_0) \approx (0.26, 0.41, 0.33)^{\mathrm{T}}$.

f) The steady state $\mathbf{s} = \lim_{n\to\infty} \mathbf{p}(t_n)$ of the system is given by

$$\mathbf{s} = \mathbf{M}\,\mathbf{s}$$

hence, it can be calculated by solving the homogenous equation

$$0 = (\mathbf{M} - \mathbb{I})\,\mathbf{s}$$

and using the fact that $\sum_i s_i = 1$, i.e. $\mathbf{s}$ is a probability vector; or, alternatively, compute the eigenvector for the eigenvalue problem

$$\det \mathbf{M} - \lambda \cdot \mathbb{I} = 0$$

for $\lambda = 1$.

```
In [4]: # use the fact that the abs. max. eigenvalue is 1.0
        s = np.linalg.eig(M)[1][:, np.argmax(np.absolute(np.linalg.eig(M)[0]))]
        # rescale
        s = (1./(np.sum(s))) * s
        print('steady state: ', s)
```

```
steady state:  [ 0.265625+0.j  0.406250+0.j  0.328125+0.j]
```

Hence, the (rounded) steady state is given by

$$\mathbf{s} \approx \begin{pmatrix} 0.26 \\ 0.41 \\ 0.33 \end{pmatrix} = \begin{pmatrix} s_r \\ s_c \\ s_s \end{pmatrix}$$

in accordance with exercise *e)* and such that $\sum_i s_i = 1$.