

SML - Übung 1



Task 1:

1a) : $\underline{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \dots & & a_{nn} \end{bmatrix}$ $\underline{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & & & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & \dots & & b_{nn} \end{bmatrix}$ $\underline{C} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & & & c_{2n} \\ \vdots & & & \vdots \\ c_{n1} & \dots & & c_{nn} \end{bmatrix}$

① commutative : $(\underline{A} \cdot \underline{B}) \cdot \underline{C} = \begin{bmatrix} \sum_{i=1}^n a_{i1} \cdot b_{i1} & \dots & \sum_{i=1}^n a_{in} \cdot b_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{ni} \cdot b_{i1} & \dots & \sum_{i=1}^n a_{ni} \cdot b_{in} \end{bmatrix} \cdot \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} \sum_{j=1}^n \left(\sum_{i=1}^n a_{i1} \cdot b_{ij} \right) \cdot c_{j1} & \dots & \sum_{j=1}^n \left(\sum_{i=1}^n a_{in} \cdot b_{ij} \right) \cdot c_{jn} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^n \left(\sum_{i=1}^n a_{ni} \cdot b_{ij} \right) \cdot c_{j1} & \dots & \sum_{j=1}^n \left(\sum_{i=1}^n a_{ni} \cdot b_{ij} \right) \cdot c_{jn} \end{bmatrix}$$

$$\underline{A} \cdot (\underline{B} \cdot \underline{C}) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} \sum_{i=1}^n b_{i1} \cdot c_{i1} & \dots & \sum_{i=1}^n b_{in} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n b_{ni} \cdot c_{i1} & \dots & \sum_{i=1}^n b_{ni} \cdot c_{in} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^n a_{1j} \cdot \left(\sum_{i=1}^n b_{ji} \cdot c_{i1} \right) & \dots & \sum_{j=1}^n a_{1j} \cdot \left(\sum_{i=1}^n b_{ji} \cdot c_{in} \right) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^n a_{nj} \cdot \left(\sum_{i=1}^n b_{ji} \cdot c_{i1} \right) & \dots & \sum_{j=1}^n a_{nj} \cdot \left(\sum_{i=1}^n b_{ji} \cdot c_{in} \right) \end{bmatrix}$$

element analys: $\because \sum_{j=1}^n \left(\sum_{i=1}^n a_{i1} \cdot b_{ij} \right) \cdot c_{j1} \neq \sum_{j=1}^n a_{1j} \cdot \left(\sum_{i=1}^n b_{ji} \cdot c_{i1} \right)$

$\therefore (\underline{A} \cdot \underline{B}) \cdot \underline{C} \neq \underline{A} \cdot (\underline{B} \cdot \underline{C})$

\Rightarrow commutative disproved

② distributive: $(\underline{A} + \underline{B}) \cdot \underline{C} = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nn} + b_{nn} \end{bmatrix} \cdot \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} \sum_{i=1}^n (a_{i1} + b_{i1}) \cdot c_{i1} & \dots & \sum_{i=1}^n (a_{in} + b_{in}) \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n (a_{ni} + b_{ni}) \cdot c_{i1} & \dots & \sum_{i=1}^n (a_{ni} + b_{ni}) \cdot c_{in} \end{bmatrix}$$

associative

$$= \begin{bmatrix} \sum_{i=1}^n a_{i1} \cdot c_{i1} & \dots & \sum_{i=1}^n a_{in} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{ni} \cdot c_{i1} & \dots & \sum_{i=1}^n a_{ni} \cdot c_{in} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^n b_{i1} \cdot c_{i1} & \dots & \sum_{i=1}^n b_{in} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n b_{ni} \cdot c_{i1} & \dots & \sum_{i=1}^n b_{ni} \cdot c_{in} \end{bmatrix}$$

$= \underline{A} \cdot \underline{C} + \underline{B} \cdot \underline{C} \Rightarrow$ distributive proved

③ associative: $(\underline{A} + \underline{B}) + \underline{C} = \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \dots & a_{nn}+b_{nn} \end{bmatrix} + \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} a_{11}+b_{11}+c_{11} & \dots & a_{1n}+b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1}+c_{n1} & \dots & a_{nn}+b_{nn}+c_{nn} \end{bmatrix}$$

$$\underline{A} + (\underline{B} + \underline{C}) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} + \begin{bmatrix} b_{11}+c_{11} & \dots & b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1}+c_{n1} & \dots & b_{nn}+c_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11}+c_{11} & \dots & a_{1n}+b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1}+c_{n1} & \dots & a_{nn}+b_{nn}+c_{nn} \end{bmatrix}$$

$$\Rightarrow (\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C})$$

\Rightarrow associative proved

1b) ① method 1: elementary transformation:

$$[\underline{A} \quad \underline{I}] \xrightarrow{\text{elementary transformation}} [\underline{I} \quad \underline{A}^{-1}]$$

$$= \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 4 & 6 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

② methode 2: use adjoint matrix

$$\underline{A}^{-1} = \frac{1}{|\underline{A}|} \underbrace{\begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{bmatrix}}_{\text{adjoint}} \quad \text{wherein } A_{ij} \text{ is algebraic cofactor}$$

$$= \frac{1}{\det(\underline{A})} \cdot \begin{bmatrix} 4 & 2 & 0 \\ -1 & 1 & -3 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{with } \det(\underline{A}) = |\underline{A}| = 2$$

If A is changed to $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 6 \\ 1 & 0 & 0 \end{bmatrix}$

$\det(\underline{A}) = 0 \Rightarrow A$ is a singular matrix
 \Rightarrow irreversible

1c) for $A \in \mathbb{R}^{n \times m}$

⊕ left Moore-Penrose pseudoinverse: $\underline{A}^+_{\text{left}} = (\underline{A}^T \cdot \underline{A})^{-1} \cdot \underline{A}^T$
right Moore-Penrose pseudoinverse: $\underline{A}^+_{\text{right}} = \underline{A}^T \cdot (\underline{A} \cdot \underline{A}^T)^{-1}$

⊗ dimensional analysis
if $A \in \mathbb{R}^{2 \times 3}$

$$\underline{A}^+_{\text{left}} = (\underline{A}^T \cdot \underline{A})^{-1} \cdot \underline{A}^T$$

$$\rightarrow (3 \times 2) \cdot (2 \times 3) \cdot (3 \times 2) \rightarrow (3 \times 2)$$

$$\underline{A}^+_{\text{right}} = \underline{A}^T \cdot (\underline{A} \cdot \underline{A}^T)^{-1}$$

$$\rightarrow (3 \times 2) \cdot [(2 \times 3) \cdot (3 \times 2)] \rightarrow (3 \times 2)$$

$$\Rightarrow \underline{A}^+ \in \mathbb{R}^{3 \times 2}$$

1d): eigenvector: represent the orientation of each principal components

eigenvalue: represent the weight of each principal components

both of them represent the components distribution of all features

(PCA - perspective)

Task 2

2a):

i): for Ω is a finite set

① Expectation: $E[f(\omega)] = \sum_{\omega \in \Omega} f(\omega) \cdot P(\omega)$

$$\text{Variance: } V[f(\omega)] = \sum_{\omega \in \Omega} P(\omega) \cdot (f(\omega) - E[f(\omega)])^2$$

$$= E[(f(\omega) - E[f(\omega)])^2]$$

② Expectation and Variance are linear operators

Definition and first consequences [\[edit\]](#)

Let V and W be vector spaces over the same field K . A function $f: V \rightarrow W$ is said to be a *linear map* if for any two vectors $u, v \in V$ and any scalar $c \in K$ the following two conditions are satisfied:

$$f(u + v) = f(u) + f(v) \quad \text{additivity / operation of addition}$$

$$f(cu) = cf(u) \quad \text{homogeneity of degree 1 / operation of scalar multiplication}$$

For Expectation: the rules of expectation prove it:

$$E[aX] = a$$

$$E[X+Y] = E[X] + E[Y]$$

$$E[\sum_i a_i X_i] = \sum_i a_i E[X_i]$$

Variance is essential

$$V[f(\omega)] = E$$

	1	2	3	4	5	6
A	4	1	6	2	1	4
B	5	6	1	1	4	1
C	3	3	4	2	3	3

$$= E[(f(\omega))^2] - 2E[f(\omega)] \cdot E[f(\omega)] + (E[f(\omega)])^2$$

$$= E[(f(\omega))^2] - E[f(\omega)]^2$$

$$= \sum_{\omega \in \Omega} P(\omega) \cdot (f(\omega))^2 - E[f(\omega)]^2$$

2)

	1	2	3	4	5	6
A	4	1	6	2	1	4
B	5	6	1	1	4	1
C	3	3	4	2	3	3

X \	1	2	3	4	5	6
P(X A)	2/9	1/18	1/3	1/9	1/18	2/9
P(X B)	5/18	1/3	1/18	1/18	2/9	1/18
P(X C)	1/6	1/6	2/9	1/9	1/6	1/6

For dice A, set $P(A)=1 \Rightarrow P(X,A) = P(X|A) \cdot P(A) = P(X|A) = P_A(X)$

$$\bar{X}_A = \sum_{i=1}^6 X_i \cdot P(X_i, A) = 1 \cdot \frac{2}{9} + 2 \cdot \frac{1}{18} + 3 \cdot \frac{1}{3} + 4 \cdot \frac{1}{9} + 5 \cdot \frac{1}{18} + 6 \cdot \frac{2}{9}$$

$$= \frac{61}{18} \approx 3,3889$$

unbiased $V[X, A] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_A)^2$ (Bessel's Correction)

$$= \frac{1}{12} \cdot \left(4 \cdot \left(1 - \frac{61}{18}\right)^2 + 1 \cdot \left(2 - \frac{61}{18}\right)^2 + 6 \cdot \left(3 - \frac{61}{18}\right)^2 + 2 \cdot \left(4 - \frac{61}{18}\right)^2 + 1 \cdot \left(5 - \frac{61}{18}\right)^2 + 4 \cdot \left(6 - \frac{61}{18}\right)^2 \right)$$

$$= \frac{1013}{306} \approx 3,3105$$

For dice B, set $P(B)=1 \Rightarrow P(X,B) = P(X|B) \cdot P(B) = P(X|B) = P_B(X)$

$$\bar{X}_B = \sum_{i=1}^6 X_i \cdot P(X_i, B) = 1 \cdot \frac{5}{18} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{18} + 4 \cdot \frac{1}{18} + 5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{18}$$

$$= \frac{25}{9} \approx 2,7778$$

unbiased $V[X, B] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_B)^2$

$$= \frac{1}{12} \left(5 \cdot \left(1 - \frac{25}{9}\right)^2 + 6 \cdot \left(2 - \frac{25}{9}\right)^2 + 1 \cdot \left(3 - \frac{25}{9}\right)^2 + 1 \cdot \left(4 - \frac{25}{9}\right)^2 + 4 \cdot \left(5 - \frac{25}{9}\right)^2 + 1 \cdot \left(6 - \frac{25}{9}\right)^2 \right)$$

$$= \frac{460}{153} \approx 3,0065$$

For dice C, set $P(C)=1 \Rightarrow P(X,C) = P(X|C) \cdot P(C) = P(X|C) = P_C(X)$

$$\bar{X}_C = \sum_{i=1}^6 X_i \cdot P(X_i, C) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{1}{9} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$$

$$= \frac{31}{9} \approx 3,4444$$

unbiased $V[X, C] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_C)^2$

$$= \frac{1}{12} \cdot \left(3 \cdot \left(1 - \frac{31}{9}\right)^2 + 3 \cdot \left(2 - \frac{31}{9}\right)^2 + 4 \cdot \left(3 - \frac{31}{9}\right)^2 + 2 \cdot \left(4 - \frac{31}{9}\right)^2 + 3 \cdot \left(5 - \frac{31}{9}\right)^2 + 3 \cdot \left(6 - \frac{31}{9}\right)^2 \right)$$

$$= \frac{472}{153} \approx 3,0850$$

Dice	Expectation	Variance
A	3,3889	3,3105
B	2,7778	3,0065
C	3,4444	3,0850

$$3) \text{ KL-divergence} = - \sum_x P_i(x) \cdot \ln \frac{Q(x)}{P_i(x)} = \sum_x P_i(x) \cdot \ln \frac{P_i(x)}{Q(x)} \quad i = A, B \text{ or } C$$

for unbiased distribution $Q(x)$, $\mu = 3,5$ $\sigma = \frac{3,5}{\sqrt{2}}$

$$\Rightarrow Q(x) = \frac{1}{6} \quad \text{for } x = 1, 2, \dots, 6$$

$$\text{Dice A: } KL = \sum_{x=1}^6 P_A(x) \cdot \ln \frac{P_A(x)}{Q(x)}$$

$$= \frac{2}{9} \cdot \ln\left(\frac{2/9}{1/6}\right) + \frac{1}{18} \cdot \ln\left(\frac{1/18}{1/6}\right) + \frac{1}{3} \cdot \ln\left(\frac{1/3}{1/6}\right) + \frac{1}{9} \cdot \ln\left(\frac{1/9}{1/6}\right) + \frac{1}{18} \cdot \ln\left(\frac{1/18}{1/6}\right) + \frac{2}{9} \cdot \ln\left(\frac{2/9}{1/6}\right)$$

$$= 0.1918$$

$$\text{Dice B: } KL = \sum_{x=1}^6 P_B(x) \cdot \ln \frac{P_B(x)}{Q(x)}$$

$$= \frac{5}{18} \cdot \ln\left(\frac{5/18}{1/6}\right) + \frac{1}{3} \cdot \ln\left(\frac{1/3}{1/6}\right) + \frac{1}{18} \cdot \ln\left(\frac{1/18}{1/6}\right) + \frac{1}{18} \cdot \ln\left(\frac{1/18}{1/6}\right) + \frac{2}{9} \cdot \ln\left(\frac{2/9}{1/6}\right) + \frac{1}{18} \cdot \ln\left(\frac{1/18}{1/6}\right)$$

$$= 0.2538$$

$$\text{Dice C: } KL = \sum_{x=1}^6 P_C(x) \cdot \ln \frac{P_C(x)}{Q(x)}$$

$$= \frac{1}{6} \cdot \ln\left(\frac{1/6}{1/6}\right) + \frac{1}{6} \cdot \ln\left(\frac{1/6}{1/6}\right) + \frac{2}{9} \cdot \ln\left(\frac{2/9}{1/6}\right) + \frac{1}{9} \cdot \ln\left(\frac{1/9}{1/6}\right) + \frac{1}{6} \cdot \ln\left(\frac{1/6}{1/6}\right) + \frac{1}{6} \cdot \ln\left(\frac{1/6}{1/6}\right)$$

$$= 0.0189$$

2b)

1) $A : \{ \text{a Person has a cold} \}$ $B : \{ \text{a person has back-pain.} \}$

2) $C : \{ \text{World population} \}$ $D : \{ \text{healthy people, who do not have cold} \}$

3) $\odot \{ \text{A person with a cold has back-pain 30\% of the time} \}$

$$\Rightarrow P(B|A) = 30\%$$

$\odot \{ \text{3\% of the world population has a cold} \}$

$$\Rightarrow P(A|C) = 3\% \quad \text{mit } P(C) = 1 \Rightarrow P(A) = 3\%$$

$\odot \{ \text{10\% of those who do not have a cold still have back-pain} \}$

$$\Rightarrow P(B|D) = 10\% \quad \text{mit } P(D) = 1 - P(A|C) = 97\%$$

$$4) P(B) = P(B,A) + P(B,D)$$

$$= P(B|A) \cdot P(A) + P(B|D) \cdot P(D)$$

$$= 30\% \cdot 3\% + 10\% \cdot 97\% = 10.6\%$$

Bayes' theorem

$$\Downarrow$$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

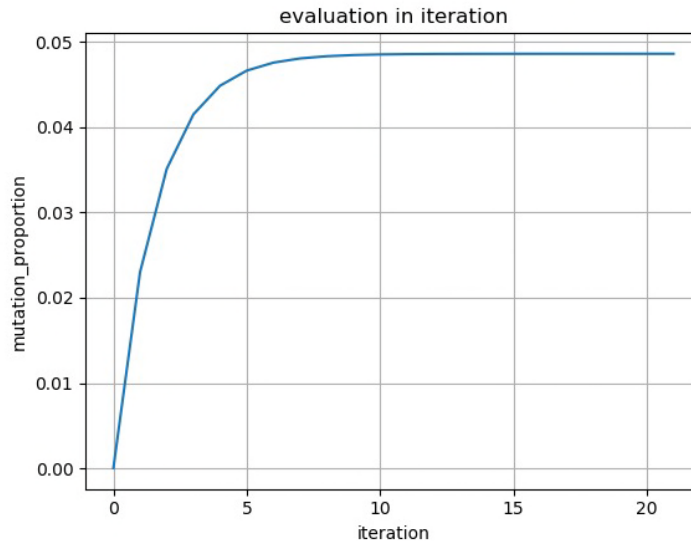
$$= \frac{30\% \cdot 3\%}{10.6\%} = 8.49\%$$

2c) Markov chain

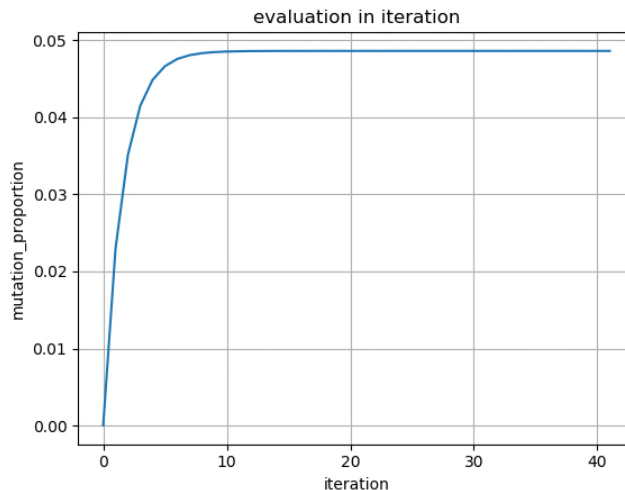
1): $\vec{s}_t = \begin{pmatrix} m \\ \tilde{m} \end{pmatrix}$ with probability matrix: $\underline{I} = \begin{bmatrix} 55\% & 2,3\% \\ 45\% & 97,7\% \end{bmatrix}$

$$\Rightarrow \vec{s}_{t+1} = \underline{I} \cdot \vec{s}_t = \begin{bmatrix} 55\% & 2,3\% \\ 45\% & 97,7\% \end{bmatrix} \begin{bmatrix} m \\ \tilde{m} \end{bmatrix} = \begin{pmatrix} 55\% m + 2,3\% \tilde{m} \\ 45\% m + 97,7\% \tilde{m} \end{pmatrix}$$

2)



3)



approximate at 7 step will culture contain stop to change mutation

if $55\% m + 2,3\% \tilde{m} = m$ mit $\tilde{m} = (1-m)$ $\Rightarrow m \approx 4,86\%$

\Rightarrow Evaluation process get stable probability

