ML2 Home Assignment 2

id: 12179078

Volodymyr Medentsiy

№1.

1. $I(X;Y) = KL(p(x,y) \| p(x)\cdot p(y)) = H(X) - H(X|Y)$

– measures how we reduce uncertainty of x if we observe y.

$I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$

– measures how much information on average Y convey about X given known Z.

2. $I(X,Y) = \sum_{x,y} p(x,y) \log \frac{p_{x,y}(x,y)}{p_x(x)\, p_y(y)}$

Let us find $p(x)$, $p(y)$ and $p(x,y)$:

x	$p(x)$
0	$0{,}192 + 0{,}144 + 0{,}048 + 0{,}216 = 0{,}6$
1	$0{,}192 + 0{,}064 + 0{,}048 + 0{,}096 = 0{,}4$

y	$p(y)$
0	$0{,}192 + 0{,}144 + 0{,}192 + 0{,}064 = 0{,}592$
1	$0{,}048 + 0{,}216 + 0{,}048 + 0{,}096 = 0{,}408$

x	y	$p(x,y)$
0	0	$0{,}192 + 0{,}144 = 0{,}336$
0	1	$0{,}216 + 0{,}048 = 0{,}264$
1	0	$0{,}192 + 0{,}064 = 0{,}256$
1	1	$0{,}048 + 0{,}096 = 0{,}144$

Then $I(X;Y) = 0{,}336 \cdot \log \frac{0{,}336}{0{,}6 \cdot 0{,}592} + 0{,}264 \log \frac{0{,}264}{0{,}6 \cdot 0{,}408} +$

$+ 0{,}256 \log \frac{0{,}256}{0{,}4 \cdot 0{,}592} + 0{,}144 \log \frac{0{,}144}{0{,}4 \cdot 0{,}408} \approx 0{,}0032$

$i(X;Y) > 0$, which means that X and Y are conditionally dependent and observing Y reduces uncertainty in X.

3. $i(X;Y|Z) = \sum_{z} p(z)\, D_{KL}\left(p(X,Y|z) \,\|\, p(X|z)\,p(Y|z)\right) =$

$$= \sum_{z} p(z) \cdot \left(-\sum_{x,y} p(x,y|z) \log \frac{p(x|z)\,p(y|z)}{p(x,y|z)}\right)$$

Let us find $p(z)$, $p(x|z)$, $p(y|z)$, $p(x,y|z)$:

z	$p(z)$
0	$0{,}192 + 0{,}048 + 0{,}192 + 0{,}048 = 0{,}48$
1	$0{,}144 + 0{,}216 + 0{,}064 + 0{,}096 = 0{,}52$

$x \mid z=0$	$p(x \mid z=0) = \frac{p(x, z=0)}{p(z=0)}$
0	$(0{,}192 + 0{,}048)/0{,}48 = 1/2$
1	$(0{,}192 + 0{,}048)/0{,}48 = 1/2$

$x \mid z=1$	$p(x \mid z=1) = \frac{p(x, z=1)}{p(z=1)}$
0	$(0{,}144 + 0{,}216)/0{,}52 \approx 0{,}692$
1	$(0{,}064 + 0{,}096)/0{,}52 \approx 0{,}308$

$y \mid z=0$	$p(y \mid z=0)$
0	$0{,}192 \cdot 2/0{,}48 = 0{,}8$
1	$0{,}048 \cdot 2/0{,}48 = 0{,}2$

$y \mid z=1$	$p(y \mid z=1)$
0	$(0{,}144 + 0{,}064)/0{,}52 = 0{,}4$
1	$(0{,}216 + 0{,}096)/0{,}52 = 0{,}6$

$x, y \mid z=0$	$p(x,y \mid z=0)$
0 0	$0{,}192 / 0{,}48 = 0{,}4$
0 1	$0{,}048 / 0{,}48 = 0{,}1$
1 0	$0{,}192 / 0{,}48 = 0{,}4$
1 1	$0{,}048 / 0{,}48 = 0{,}1$

$x, y \mid z=1$	$p(x,y \mid z=1)$
0 0	$0{,}144 / 0{,}52 = 0{,}277$
0 1	$0{,}216 / 0{,}52 = 0{,}415$
1 0	$0{,}064 / 0{,}52 = 0{,}123$
1 1	$0{,}096 / 0{,}52 = 0{,}185$

Thus $I(X;Y|Z) = 0{,}48 \cdot \left(-\left(0{,}4 \log \frac{0{,}5 \cdot 0{,}8}{0{,}4} + 0{,}1 \log \frac{0{,}5 \cdot 0{,}2}{0{,}1} +\right.\right.$

$\left.\left. + 0{,}4 \log \frac{0{,}5 \cdot 0{,}8}{0{,}4} + 0{,}1 \log \frac{0{,}5 \cdot 0{,}2}{0{,}1}\right)\right) +$

$+ 0{,}52 \cdot \left(-\left(0{,}277 \log \frac{0{,}692 \cdot 0{,}4}{0{,}277} + 0{,}415 \log \frac{0{,}692 \cdot 0{,}6}{0{,}415} +\right.\right.$

$\left.\left. + 0{,}123 \log \frac{0{,}308 \cdot 0{,}4}{0{,}123} + 0{,}185 \log \frac{0{,}308 \cdot 0{,}6}{0{,}185}\right)\right) = 0.$

4. Let us find $p(z|x)$

$z \mid x=0$	$p(z \mid x=0)$
0	$(0{,}192 + 0{,}048)/0{,}6 = 0{,}4$
1	$(0{,}144 + 0{,}216)/0{,}6 = 0{,}6$

$z \mid x=1$	$p(z \mid x=1)$
0	$(0{,}192 + 0{,}048)/0{,}4 = 0{,}6$
1	$(0{,}096 + 0{,}064)/0{,}4 = 0{,}4$

To show that $p(x,y,z) = p(x)p(z|x)p(y|z)$ I construct the table:

x	y	z	$p(x,y,z)$	$p(x)$	$p(z\mid x)$	$p(y\mid z)$	$p(x)p(z\mid x)p(y\mid z)$
0	0	0	0,192	0,6	0,4	0,8	0,192
0	0	1	0,144	0,6	0,6	0,4	0,144
0	1	0	0,048	0,6	0,4	0,2	0,048
0	1	1	0,216	0,6	0,6	0,6	0,216
1	0	0	0,192	0,4	0,6	0,8	0,192
1	0	1	0,064	0,4	0,4	0,4	0,064
1	1	0	0,048	0,4	0,6	0,2	0,048
1	1	1	0,096	0,4	0,4	0,6	0,096

Corresponding DAG:

$$X \longrightarrow Z \longrightarrow Y$$

№3

1. $p(x) = N(x|\mu, \Sigma)$, $q(x) = N(x|m, L)$. $x \in \mathbb{R}^K$

$$KL(p\|q) = -\int p(x) \ln \frac{q(x)}{p(x)} dx = -\int \frac{1}{(2\pi)^{K/2}\sqrt{\det\Sigma}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \cdot$$

$$\cdot \ln \frac{\sqrt{\det\Sigma}\, e^{-\frac{1}{2}(x-m)^T L^{-1}(x-m)}}{\sqrt{\det L}\, e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}} dx = -\left(\int p(x) \ln\sqrt{\frac{\det\Sigma}{\det L}}\, dx + \right.$$

$$\left. + \int p(x) \left(-\frac{1}{2}(x-m)^T L^{-1}(x-m) + \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) dx \right) =$$

$$= -\left(\frac{1}{2} \ln \frac{\det\Sigma}{\det L} - \frac{1}{2}\int p(x)(x-m)^T L^{-1}(x-m)\, dx + \right.$$

$$\left. + \frac{1}{2}\int p(x)(x-\mu)^T\Sigma^{-1}(x-\mu)\, dx \right) = -\frac{1}{2} \ln\frac{\det\Sigma}{\det L} + \frac{1}{2}\Big((\mu-m)^T L^{-1}(\mu-m) +$$

$$+ \mathrm{Tr}(L^{-1}\Sigma)\Big) - \frac{1}{2}\left((\mu-\mu)^T\Sigma^{-1}(\mu-\mu) + \mathrm{Tr}(\Sigma^{-1}\Sigma)\right) =$$

$$= -\frac{1}{2}\ln\frac{\det\Sigma}{\det L} + \frac{1}{2}\left((\mu-m)^T L^{-1}(\mu-m) + \mathrm{Tr}(L^{-1}\Sigma)\right) - \frac{K}{2}.$$

2) $H(p) = -\int p(x) \ln p(x)\, dx =$

$$= -\int \frac{1}{(2\pi)^{K/2}\sqrt{\det\Sigma}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \ln \frac{1}{(2\pi)^{K/2}\sqrt{\det\Sigma}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} dx =$$

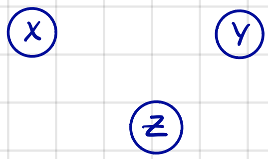
$$= \frac{1}{2}\int p(x)(x-\mu)^T\Sigma^{-1}(x-\mu)\, dx + \int p(x) \ln\left((2\pi)^{K/2}\sqrt{\det\Sigma}\right) dx =$$

$$= \frac{1}{2}\cdot\left((\mu-\mu)^T\Sigma^{-1}(\mu-\mu) + \mathrm{Tr}(\Sigma^{-1}\Sigma)\right) + \frac{1}{2}\ln\left((2\pi)^K \det\Sigma\right) =$$

$$= \frac{K}{2} + \frac{K}{2}\ln 2\pi + \frac{1}{2}\ln\det\Sigma.$$

W2.

I

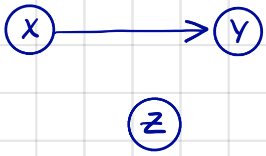


$X \perp\!\!\!\perp Y \mid *$

$Z \perp\!\!\!\perp X \mid *$

$Y \perp\!\!\!\perp Z \mid *$

II

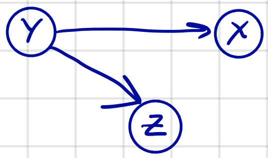


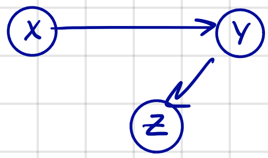
$X \not\perp\!\!\!\perp Y \mid *$

$X \perp\!\!\!\perp Z \mid *$

$Y \perp\!\!\!\perp Z \mid *$

III





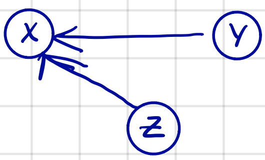
$X \perp\!\!\!\perp Z \mid Y$

$X \not\perp\!\!\!\perp Y \mid *$

$Y \perp\!\!\!\perp Z \mid *$

$X \perp\!\!\!\perp Z \mid *$

IV



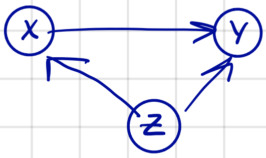
$Y \perp\!\!\!\perp Z \mid *$

$X \not\perp\!\!\!\perp Y \mid *$

$X \not\perp\!\!\!\perp Z \mid *$

$Y \not\perp\!\!\!\perp Z \mid X$

V



$X \not\perp\!\!\!\perp Y \mid *$

$Y \not\perp\!\!\!\perp Z \mid *$

$X \not\perp\!\!\!\perp Z \mid *$

There are 5 clusters which are same up to permutations.