Machine Learning 2

Exercise Sheet 3

Exercise 1: Maximum Entropy Distributions

(a)

$$\begin{aligned}\mathcal{L}\left((s(x)_{x\in\mathcal{G}}),\lambda_1,\lambda_2,\lambda_3\right) = &-\sum_{x\in G}\exp(s(x))s(x)\delta + \lambda_1\left(\sum_{x\in\mathcal{G}}\exp(s(x))\delta - 1\right)\\ &+\lambda_2\left(\sum_{x\in\mathcal{G}}x\cdot\exp(s(x))\delta - 0\right)\\ &+\lambda_3\left(\sum_{x\in\mathcal{G}}x^2\exp(s(x))\delta - \sigma^2\right)\end{aligned}$$

(b)

$$\frac{\partial\mathcal{L}}{\partial s(x)} = -\exp(s(x))(s(x)+1) + \lambda_1\exp(s(x)) + \lambda_2 x\exp(s(x)) + \lambda_3 x^2$$

Setting this to zero leads to

$$-s(x) - 1 + \lambda_1 + \lambda_2 x + \lambda_3 x^2 = 0$$

and thus,

$$p(x) = \exp(s(x)) = \exp(\lambda_3 x^2 + \lambda_2 x + \lambda_1 - 1).$$

Exercise 2: Independent Components in Two Dimensions

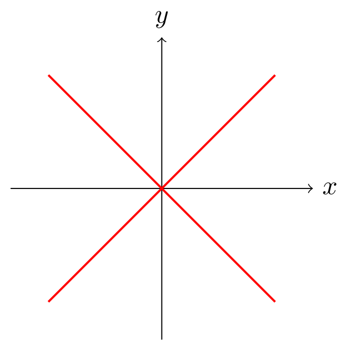
(a)

$$\omega = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$$

$$\|\omega\|^2 = \cos^2\theta + \sin^2\theta = 1.$$

$$\int\delta(x)\mathrm{d}x = 1$$

$$\int y\delta(x-y)\mathrm{d}x = x$$



(b) We have

$$
\begin{aligned}
\mathrm{E}\left[z\right] &= \int\int z(x,y)\cdot p(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \int\int (x\cos\theta + y\cos\theta)\cdot p(x)\left(\frac{1}{2}\delta(y-x)+\frac{1}{2}\delta(y+x)\right)\mathrm{d}x\mathrm{d}y \\
&= \int\left(\frac{1}{2}\left(x\cos\theta + x\sin\theta\right)+\frac{1}{2}\left(x\cos\theta - x\sin\theta\right)\right)p(x)\,\mathrm{d}x \\
&= \left(\frac{\cos\theta+\sin\theta}{2}+\frac{\cos\theta-\sin\theta}{2}\right)\underbrace{\int x\,p(x)\,\mathrm{d}x}_{\mathrm{E}[x]=0} \\
&= 0
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}\left[z^2\right] &= \int\left(\frac{1}{2}\left(x\cos\theta + x\sin\theta\right)^2+\frac{1}{2}\left(x\cos\theta - x\sin\theta\right)^2\right)\,p(x)\,\mathrm{d}x \\
&= \left(\frac{(\cos\theta+\sin\theta)^2}{2}+\frac{(\cos\theta-\sin\theta)^2}{2}\right)\underbrace{\int x^2\,p(x)\,\mathrm{d}x}_{\mathrm{E}[x^2]=\mathrm{Var}[x]=1} \\
&= \cos^2\theta+\sin^2\theta \\
&= 1.
\end{aligned}
$$

Since the variance is the same in any direction, there are no principal components.

(c)

$$
\begin{aligned}
\mathrm{E}\left[z^4\right] - 3 &= \left(\frac{(\cos\theta + \sin\theta)^4}{2} + \frac{(\cos\theta + \sin\theta)^4}{2} \right) \underbrace{\int x^4 \; p(x) \, \mathrm{d}x}_{\mathrm{E}[x^4]=3} - 3 \\
&= \cos^4\theta + \sin^4\theta + 2\cos^2\theta\sin^2\theta + 4\cos^2\theta\sin^2\theta \\
&= \left(\cos^2\theta + \sin^2\theta\right)^2 + 4\left(\cos\theta\sin\theta\right)^2 \\
&= 1 + \left(\sin(2\theta)\right)^2 \\
&= \frac{1 - \cos(4\theta)}{2}
\end{aligned}
$$

Exercise 3: Deriving a Special Case of FastICA

(a)

(b)

(c)

(d)