Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Machine Learning 1

Winter semester 2016/17

Group APXNLE

Exercise 1

Members

Jing Li[387272] jing.li.1@campus.tu-berlin.de

Kumar Awanish[386697] k.awanish@campus.tu-berlin.de

Manjiao Xu[386498] manjiao.xu@campus.tu-berlin.de

Rudresha Gulaganjihalli Parameshappa[386642]

Gulaganjihalliparameshappa@ campus.tu-berlin.de

Sonali Nayak[386995] sonali.nayak@campus.tu-berlin.de

Maximilian Ernst[364862] maximilian.ernst@campus.tu-berlin.de

Exercise 1: Estimating the Bayes Error

(a)

1a) Assume: $P(\omega_1|x) < P(\omega_2|x)$

$$\to \frac{2}{\frac{1}{P(\omega_1|x)} + \frac{1}{P(\omega_2|x)}} > \frac{2}{\frac{1}{P(\omega_1|x)} + \frac{1}{P(\omega_1|x)}} = P(\omega_1|x)$$

$P(\omega_2|x) \leq P(\omega_1|x)$ analogous.

(b)

$$P(\text{error}) \leq \int \frac{2}{\frac{1}{P(\omega_1|x)} + \frac{1}{P(\omega_2|x)}} p(x)\,dx =$$

$$\int \frac{2}{\frac{p(x)}{P(\omega_1)P(x|\omega_1)} + \frac{p(x)}{P(\omega_2)P(x|\omega_2)}} p(x)\,dx =$$

$$= \int \frac{2}{\frac{(1+(x-\mu)^2)\pi}{P(\omega_1)} + \frac{(1+(x+\mu)^2)\pi}{P(\omega_2)}} dx =$$

$$\frac{1}{\pi}\int \frac{2P(\omega_1)P(\omega_2)}{(1+(x-\mu)^2)P(\omega_2) + (1+(x+\mu)^2)P(\omega_1)} dx$$

$$\frac{1}{\pi}\int \frac{2P(\omega_1)P(\omega_2)\,dx}{x^2(P(\omega_1)+P(\omega_2)) + 2x\mu(P(\omega_1)-P(\omega_2)) + (1+\mu^2)(P(\omega_1)+P(\omega_2))}$$

$$\frac{2}{\pi}\frac{2\pi\, P(\omega_1)P(\omega_2)}{\sqrt{4(P(\omega_1)+P(\omega_2))^2(1+\mu^2) - (2\mu)^2(P(\omega_1)-P(\omega_2))^2}}$$

$$= \frac{2P(\omega_1)P(\omega_2)}{\sqrt{P(\omega_1)^2 + (4\mu^2+2)P(\omega_1)P(\omega_2) + P(\omega_2)^2}}$$

(c)

Exercise 2: Bayes Decision Boundaries

(a)

2a) $P(\omega_1|x) = P(\omega_2|x) \Leftrightarrow$

$$\frac{p(x|\omega_1)P(\omega_1)}{p(x)} = \frac{p(x|\omega_2)P(\omega_2)}{p(x)} \Leftrightarrow$$

$$\Leftrightarrow \frac{p(x|\omega_1)}{p(x|\omega_2)} = \frac{P(\omega_2)}{P(\omega_1)}$$

$$\Leftrightarrow \exp\left(\frac{-|x-\mu| + |x+\mu|}{6}\right) = \frac{P(\omega_2)}{P(\omega_1)} \quad | \ln$$

$$\Leftrightarrow |x+\mu| - |x-\mu| = 6\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$$

1st case: $x+\mu > 0$ and $x-\mu < 0$

$$\Rightarrow x+\mu + x-\mu = 6\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right) \Leftrightarrow$$

$$\Leftrightarrow x = \frac{6}{2}\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$$

2nd case: $x+\mu < 0$ and $x-\mu > 0$:

$$\Rightarrow -(x+\mu) - (x-\mu) = 6\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$$

$$\Leftrightarrow -x-\mu-x+\mu = 6\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$$

$$\Leftrightarrow x = -\frac{6}{2}\ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$$

For the other cases, x will cancle out, and equality ho

only if $\pm 2\mu = 6\ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right)$. Those cases are

relevant.

(b)

2b) $P(\omega_1|x) > P(\omega_2|x) \iff P(\text{error}|x) = P(\omega_2|x)$ ∀

$\iff |x+\mu| - |x-\mu| > \sigma \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$ (*) $\forall x \in \mathbb{R}$

1st Case: $P(\omega_2) > P(\omega_1) \Rightarrow \sigma \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right) > 0$

~~(*) holds can only hold, if $\mu = 0$~~

$\Rightarrow$ in this case, (*) can never hold

If $P(\omega_2) < P(\omega_1)$, we set $\mu = 0$, thus we get $|x| - |x| > \sigma \ln\underbrace{\left(\frac{P(\omega_2)}{P(\omega_1)}\right)}_{<0}$

(c)

2c) As above: $P(\omega_1|x) = P(\omega_2|x) \iff$

$\iff \frac{P(x|\omega_1)}{P(x|\omega_2)} = \frac{P(\omega_2)}{P(\omega_1)}$

$\iff (x+\mu)^2 - (x-\mu)^2 = 2\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$

$\iff x^2 + 2\mu x + \mu^2 - x^2 + 2\mu x - \mu^2 = 2\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$

$\iff 4\mu x = 2\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$

$\iff \mu x = \frac{1}{2}\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$

$\Rightarrow$ decision boundary given by linear function

$f(x) := \mu x - \frac{1}{2}\sigma^2 \ln\left(\left(\frac{P(\omega_2)}{P(\omega_1)}\right)\right)$

For the second part it must hold: $\mu x > \frac{1}{2}\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$ $\forall x \in \mathbb{R}$, this inequality can never hold for all $x \in \mathbb{R}$.

If $\mu = 0 \Rightarrow 0 > \frac{1}{2}\sigma^2 \ln\left(\frac{P(\omega_2)}{P(\omega_1)}\right)$, this holds, if $P(\omega_1) > P(\omega_2)$

If $\mu < 0$ the inequality holds if and only if, $P(\omega_1) = 0$, $P(\omega_2) = 1$, if $\mu > 0$, we get $P(\omega_2) = 0$, $P(\omega_1) = 1$

Exercise 3: Programming

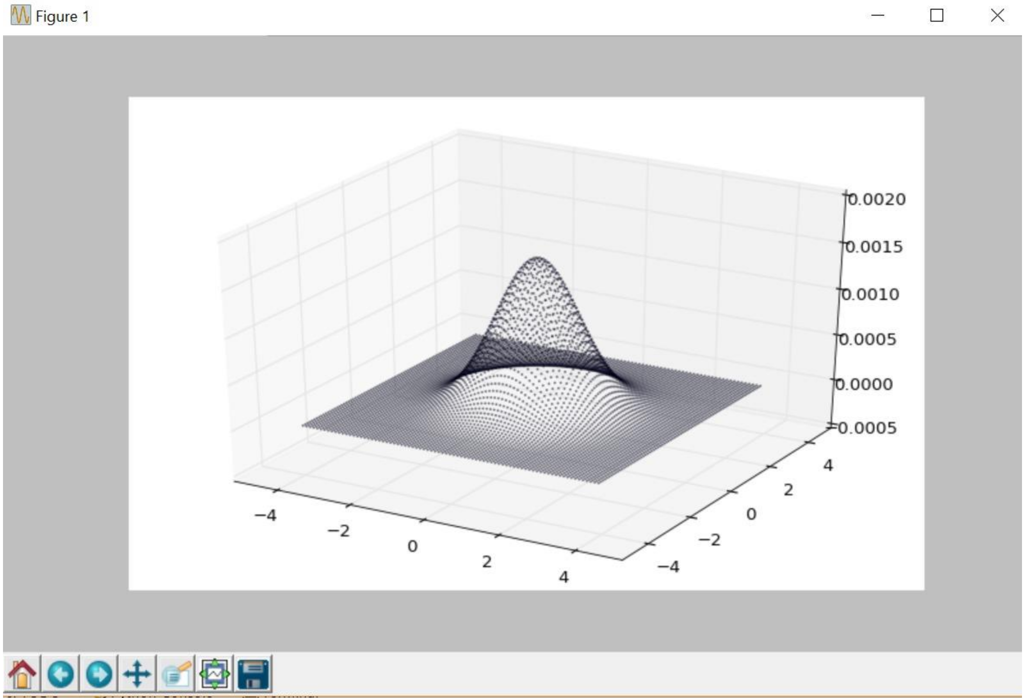
Exercise 1: Gaussian distributions

in [7]

Code:

```
import matplotlib
import numpy
import math
from matplotlib import pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
R = numpy.arange(-4,4,0.1)
X,Y = numpy.meshgrid(R,R)
print(X.shape,Y.shape)
F = numpy.sum(math.e**((-0.5)*(X**2+Y**2)))
#F=numpy._sum_(math.e**((-0.5)(X**2+Y**2)))
P=(1/F)*(math.e**((-0.5)*(X**2+Y**2)))
print(F.shape)
fig = plt.figure(figsize=(10,6))
ax = plt.axes(projection='3d')
ax.scatter(X,Y,P,s=1,alpha=0.5)
plt.show()
```

Output:



[in]8

Code:

```
import numpy as np
import math
import matplotlib.pyplot as plot
import mpl_toolkits.mplot3d.axes3d
R = np.arange(-4, 4+1e-9, 0.1)
X, Y = np.meshgrid(R, R)
Z = np.sum(math.e**(-0.5*(X**2+Y**2))) # Z = sum
(e^(-0.5*(X^2)*(Y^2)))
P = (1/Z)*math.e**(-0.5*(X**2+Y**2)) # P(x,y)
= (1/Z)*(e^(-0.5*(X^2)*(Y^2)))
# reset the peak
invalid_xy = (X**2+Y**2)<1 # thY, P, s
=0.5, alpha=0.5)
P[invalid_xy] = 0
# plot the result
fig = plot.figure(figsize=(10, 6))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(X, Y, P, s=0.5, alpha=0.5)
plot.show()
```

Output：