Pattern Recognition

Lecture 4. Review on Bayesian Decision Theory and Practice

Dr. Shanshan ZHAO

shanshan.zhao@xjtlu.edu.cn

School of AI and Advanced Computing

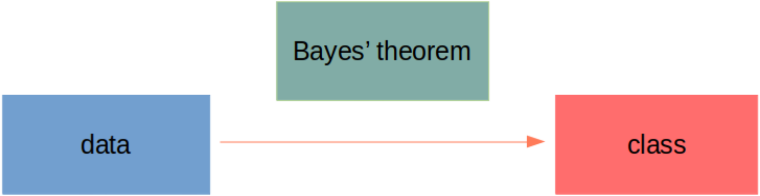
Xi'an Jiaotong-Liverpool University

Academic Year 2021-2022

Table of Contents

Classification with Bayes' theorem

In the previous lecture, we use **Bayes' theorem** for classification, to relate the probability density function of the data given the class to the posterior probability of the class given the data.



Gaussian Classifiers

When we consider the univariate Gaussian distribution, with a continuous variable x, whose pdf, given class $\omega = k$, is a Gaussian with mean μ_k and variance σ_k^2.
Using Bayes' theorem we write:

$$
\begin{aligned}
P(\omega_k|x) = \frac{p(x|\omega_k)P(\omega_k)}{p(x)} &\propto p(x|\omega_k)P(\omega_k) \\
&\propto N(x;\mu_k,\sigma_k^2)P(\omega_k) \\
&\propto \frac{1}{\sqrt{2\pi\sigma_k^2}} exp(\frac{-(x-\mu_k)^2}{2\sigma_k^2})P(\omega_k)
\end{aligned}
$$

Gaussian Classifiers

Log likelihoods and log probabilities When dealing Gaussians, it is often useful to take logs:

$$\begin{aligned}
\ln p(x|\mu_k, \sigma_k^2) &= \ln\left[\frac{1}{\sqrt{2\pi\sigma_k^2}} exp(\frac{-(x-\mu_k)^2}{2\sigma_k^2}))\right] \\
&= -\ln(\sqrt{2\pi\sigma_k^2}) - \frac{(x-\mu_k)^2}{2\sigma_k^2} \\
&= \frac{1}{2}\left(-\ln(2\pi) - \ln\sigma_k^2 - \frac{(x-\mu_k)^2}{\sigma_k^2}\right)
\end{aligned}$$

We can use Bayes' theorem to write log posterior probability $\ln P(\omega_k|x)$:

$$\begin{aligned}
\ln P(\omega_k|x) &= \ln p(x|\omega_k) + \ln P(\omega_k) + const. \\
&= \frac{1}{2}(-\ln(2\pi) - \ln\sigma_k^2 - \frac{(x-\mu_k)^2}{\sigma_k^2}) + \ln P(\omega_k) + const.
\end{aligned}$$

Gaussian Classifiers

Log probability ratio If ω_1 and ω_2 are modelled by Gaussians with means μ_1 and μ_2, and variances σ_1^2 and σ_1^2, then we can write the log odds(ratio of posterior probabilities) as follows:

$$\begin{aligned}\ln \frac{P(\omega_1|x)}{P(\omega_2|x)} &= \ln P(\omega_1|x) - \ln P(\omega_2|x) \\ &= -\frac{1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{(x-\mu_2)^2}{\sigma_2^2} + \ln \sigma_1^2 - \ln \sigma_2^2\right) \\ &+ \ln P(\omega_1) - \ln P(\omega_2).\end{aligned}$$

Example:Univariate Gaussian classifier

There is a problem with two clases, S and T. We assume that each class may be modelled by a Gaussian. The mean and the variance of each pdf are: $\mu_S = 10$, $\mu_T = 12$, $\sigma_S^2 = 1$, $\sigma_T^2 = 4$.

The following unlabelled data points are available:

$x_a = 10$, $x_b = 11$, $x_c = 6$

Question: To which class should each data point be assigned?

- (1) Assume the two classes have equal prior probabilities
- (2) Priors: $P(S) = 0.3$, $P(T) = 0.7$

YOU LOVE CODING ♡

Example: Multivariate Gaussian Classifier
We have two-dimensional data from three classes (A, B, C). The classes are assumed to have equal prior probabilities.
The training data is in files $trainA.dat$, $trainB.dat$, $trainC.dat$, test data in files $testA.dat$, $testB.dat$, $testC.dat$.
tasks:(w1d4_a.ipynb)

- 1. load and plot the data. How many data points? How many features?
- 2. get the mean and covariance of data in each class
- 3. compute the conditional probabilities of each class given the data
- 4. assign the data in a class that have the maximum posterior probability

Review of Gaussian(cont.)

Recall that when the 2-dimensional Gaussian distribution has a diagonal covariance matrix,i.e.$\Sigma = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$ In the exponent, $x^T \Sigma^{-1} x = a_{11} x_1^2 + a_{22} x_2^2$. Consider a 2-D Gaussian with mean vector of $\mu = (0, 0)^T$, and convariance matrix:

- case 1: $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- case 2: $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$
- case 3: $\Sigma = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$

Gaussian #1

Question:

1. Which case does it belong to?
2. Which type of correlation do x_1 and x_2 have,

(i) a positive correlation ; (ii) a negative correlation ; (iii) no correlation

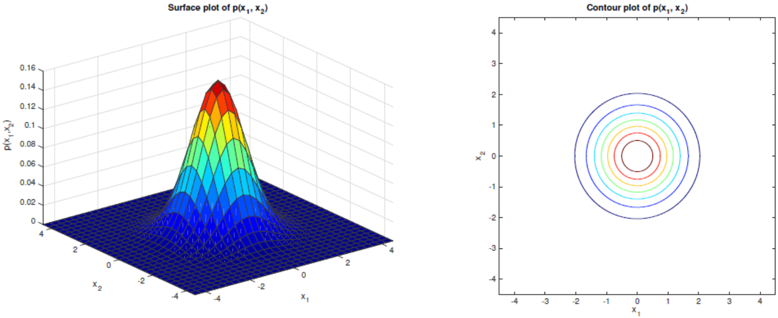


Figure: Spherical Gaussian (diagonal covariance, equal variances).

Gaussian #2

Question:

1. Which case does it belong to?
2. Which type of correlation do x_1 and x_2 have,

(i) a positive correlation ; (ii) a negative correlation ; (iii) no correlation

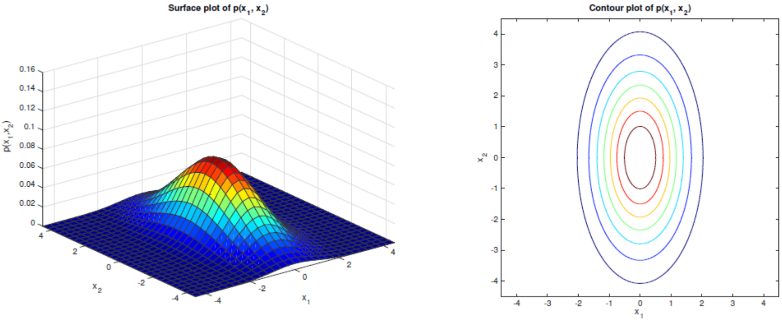


Figure: Gaussian with diagonal covariance matrix

Gaussian #3

Question:

1. Which case does it belong to?
2. Which type of correlation do x_1 and x_2 have,

(i) a positive correlation ; (ii) a negative correlation ; (iii) no correlation

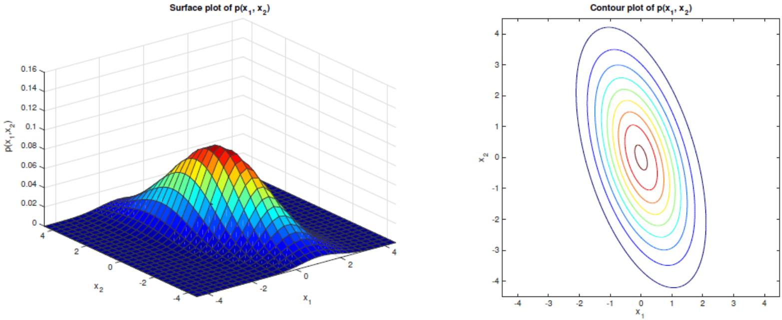


Figure: Gaussian with full covariance matrix

YOU LOVE CODING ♡

Task:(w1d4_b.ipynb)
Plot Gaussians with Python

Thank You !

$\mathcal{Q}$ & $\mathcal{A}$