

# CS509 Pattern Recognition Assignment 1

Wentao Lu (002276355), Yi Ren (002269013)

## 1 Operations with events

Let A, B and C be three independent events.

$$\begin{aligned} 1. P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B \cap C) \\ &= P(A) + P(B) + P(C) - P(A)P(B)P(C) \\ &= 1 - 0.2 \times 0.5 \times 0.3 = 0.97 \end{aligned}$$

$$2. P(A \cap B) = 0.2 = P(A)P(B) \Rightarrow P(A) = \frac{0.2}{P(B)} = \frac{0.2}{0.8} = \frac{1}{4}$$

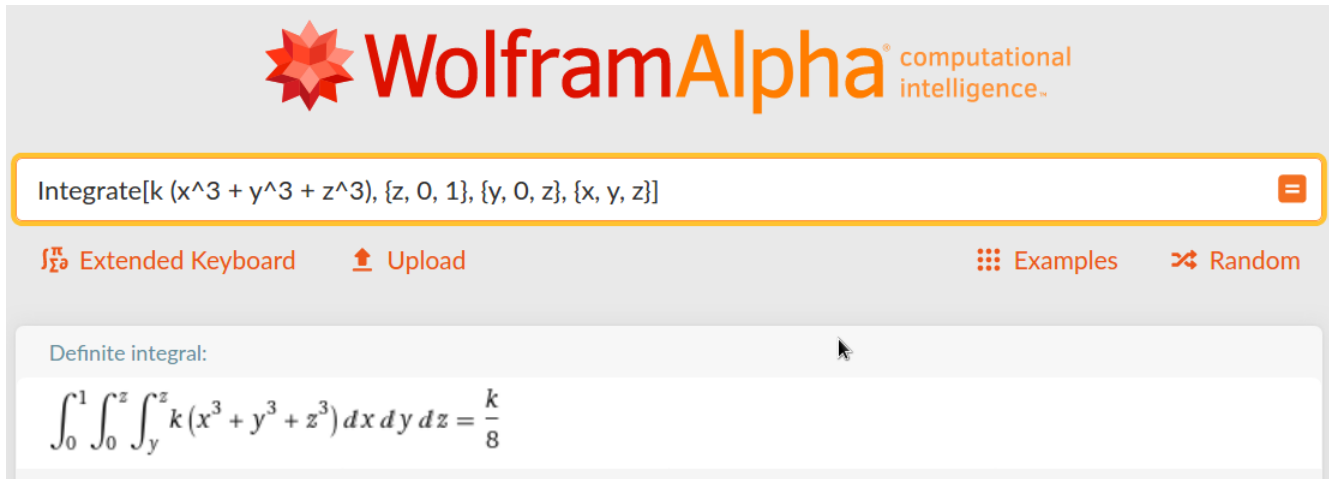
$$\begin{aligned} \Rightarrow P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A)P(B)P(C) \\ &= 0.25 + 0.8 + 0.1 - 0.25 \times 0.8 \times 0.1 \\ &= 1.15 - 0.02 = 1.13 \end{aligned}$$

$$\begin{aligned} 3. P(A) &= 1 - P(\bar{A}) = 1 - 0.2 = 0.8 \\ P(C) &= 1 - P(\bar{C}) = 1 - 0.4 = 0.6 \\ P(A \cap B) &= P(A)P(B) = 0.8P(B) = 0.16 \Rightarrow P(B) = 0.2 \end{aligned}$$

$$\begin{aligned} \Rightarrow P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A)P(B)P(C) \\ &= 0.8 + 0.2 + 0.6 - 0.8 \times 0.2 \times 0.6 \\ &= 1.504 \end{aligned}$$

## 2 Law of probability of a triplet of continuous random variables

1. Using mathematical programming languages such as WOLFRAM MATHEMATICA or SAGE, we can easily obtain that:



The screenshot shows the WolframAlpha interface. At the top is the WolframAlpha logo with the tagline "computational intelligence.". Below the logo is a search bar containing the input: `Integrate[k (x^3 + y^3 + z^3), {z, 0, 1}, {y, 0, z}, {x, y, z}]`. Below the search bar are links for "Extended Keyboard", "Upload", "Examples", and "Random". Below these links is a section titled "Definite integral:" which displays the result of the integral:  $\int_0^1 \int_0^z \int_0^z k(x^3 + y^3 + z^3) dx dy dz = \frac{k}{8}$ .

Since the integral of  $f(x, y, z)$  must sum up to 1, we have  $k/8 = 1 \Rightarrow k = 8$ . Alternatively, we can also compute integrals the hard way:

$$\begin{aligned}
 & \iiint_D (x^3 + y^3 + z^3) dx dy dz \\
 &= \iint \left[ \int_y^z (x^3 + y^3 + z^3) dx \right] dy dz = \iint \left[ \left( \frac{1}{4} x^4 + y^3 x + z^3 x \right) \Big|_{x=y}^{x=z} \right] dy dz \\
 &= \iint \left( \frac{5}{4} z^4 - \frac{5}{4} y^4 + y^3 z - z^3 y \right) dy dz \\
 &= \int \left[ \int_0^z \left( \frac{5}{4} z^4 - \frac{5}{4} y^4 + y^3 z - z^3 y \right) dy \right] dz \\
 &= \int \left[ \left( \frac{5}{4} z^4 y - \frac{y^5}{5} + \frac{z}{4} y^4 - \frac{z^3}{2} y^2 \right) \Big|_{y=0}^{y=z} \right] dz \\
 &= \int_0^1 \left( \frac{5}{4} z^5 - \frac{1}{4} z^5 + \frac{z^5}{4} - \frac{z^5}{2} \right) dz = \int_0^1 \frac{3}{4} z^5 dz \\
 &= \frac{3}{4} \cdot \frac{z^6}{6} \Big|_{z=0}^{z=1} = \frac{1}{8} = \frac{1}{k} \Rightarrow k = 8
 \end{aligned}$$

2. The marginal pdfs of  $X, Y, Z$  can be computed manually:

$$\begin{aligned}
 & \iint 8(x^3 + y^3 + z^3) dy dz \\
 &= \int \left[ \int_0^z 8(x^3 + y^3 + z^3) dy \right] dz = 8 \int \left[ (x^3 y + \frac{1}{4} y^4 + z^3 y) \Big|_{y=0}^{y=z} \right] dz \\
 &= 8 \int \left( \frac{5}{4} x^4 + z^3 x \right) dz = 8 \left( \frac{5}{4} x^4 z + x \frac{z^4}{4} \right) \Big|_{z=x}^{z=1} \\
 &= 8 \left( \frac{5}{4} x^4 + \frac{x}{4} - \frac{5}{4} x^5 - \frac{x^5}{4} \right) \\
 &= -12x^5 + 10x^4 + 2x \\
 &\Rightarrow \text{margin}(x) = \begin{cases} -12x^5 + 10x^4 + 2x, & 0 \leq x \leq 1 \\ 0, & \text{o.w.} \end{cases}
 \end{aligned}$$


---

$$\begin{aligned}
 & \iint 8(x^3 + y^3 + z^3) dx dz = 8 \int \left[ \int (x^3 + y^3 + z^3) dx \right] dz \\
 &= 8 \int \left[ \left( \frac{x^4}{4} + y^3 x + z^3 x \right) \Big|_{x=y}^{x=z} \right] dz \\
 &= 8 \int_{z=y}^{z=1} \left( \frac{5}{4} z^4 + y^3 z - \frac{5}{4} y^4 - z^3 y \right) dz \\
 &= 8 \left( \frac{z^5}{4} + y^3 \frac{z^2}{2} - \frac{5}{4} y^4 z - y \frac{z^4}{4} \right) \Big|_{z=y}^{z=1} \\
 &= 6y^5 - 10y^4 + 4y^3 - 2y + 2 \\
 &\Rightarrow \text{margin}(y) = \begin{cases} 6y^5 - 10y^4 + 4y^3 - 2y + 2, & 0 \leq y \leq 1 \\ 0, & \text{o.w.} \end{cases}
 \end{aligned}$$


---

$$\begin{aligned}
 & \iint 8(x^3 + y^3 + z^3) dx dy = 8 \int \left[ \left( \frac{x^4}{4} + y^3 x + z^3 x \right) \Big|_{x=y}^{x=z} \right] dy \\
 &= 8 \int \left( \frac{5}{4} z^4 + y^3 z - \frac{5}{4} y^4 - z^3 y \right) dy \\
 &= 8 \left( \frac{5}{4} z^4 y + z \frac{y^4}{4} - \frac{y^5}{4} - z^3 \frac{y^2}{2} \right) \Big|_{y=0}^{y=z} = 6z^5 \\
 &\Rightarrow \text{margin}(z) = \begin{cases} 6z^5, & 0 \leq z \leq 1 \\ 0, & \text{o.w.} \end{cases}
 \end{aligned}$$

However, the computation is tedious and error-prone, it'd be better to program the solution out:

Integrate[8 (x^3 + y^3 + z^3), {y, 0, x}, {z, x, 1}]

Extended Keyboard Upload Examples Random

Define integral:

$$\int_0^x \int_x^1 8(x^3 + y^3 + z^3) dz dy = 2x(-6x^4 + 5x^3 + 1)$$

Integrate[8 (x^3 + y^3 + z^3), {z, y, 1}, {x, y, z}]

Extended Keyboard Upload Examples Random

Define integral:

$$\int_y^1 \int_y^z 8(x^3 + y^3 + z^3) dx dz = 2(y-1)^2(3y^3 + y^2 + y + 1)$$

Integrate[8 (x^3 + y^3 + z^3), {y, 0, z}, {x, y, z}]

Extended Keyboard Upload Examples Random

Define integral:

$$\int_0^z \int_y^z 8(x^3 + y^3 + z^3) dx dy = 6z^5$$

As we can see, the automatic function gives the same results as we computed before.

- Since  $\text{margin}(x) * \text{margin}(y) * \text{margin}(z)$  is a very high-order polynomial expression, which apparently does not equal to  $f(x, y, z)$ , so the three variables are not independent.
- We can compute the expectation and variance using the following equations:

$$E(X) = \int_0^1 x \text{margin}(x) dx \quad (1)$$

$$E(X^2) = \int_0^1 x^2 \text{margin}(x) dx \quad (2)$$

$$\text{Var}(X) = E(X^2) - E(X)^2 \quad (3)$$

Integrate[x (-12x^5 + 10x^4 + 2x), {x, 0, 1}]

Extended Keyboard
Upload
Examples
Random

Definite integral:
More digits
Step-by-step solution

$$\int_0^1 x (-12x^5 + 10x^4 + 2x) dx = \frac{13}{21} \approx 0.61905$$

Integrate[x^2 (-12x^5 + 10x^4 + 2x), {x, 0, 1}]

Extended Keyboard
Upload
Examples
Random

Definite integral:
More digits
Step-by-step solution

$$\int_0^1 x^2 (-12x^5 + 10x^4 + 2x) dx = \frac{3}{7} \approx 0.42857$$

Hence,  $E(X) = 0.61905$ ,  $Var(X) = 0.42857 - 0.61905^2 = 0.045347097$ .

5. ...

Integrate[y (6y^5 - 10y^4 + 4y^3 - 2y + 2), {y, 0, 1}]

Extended Keyboard
Upload
Examples
Random

Definite integral:
More digits
Step-by-step solution

$$\int_0^1 y (6y^5 - 10y^4 + 4y^3 - 2y + 2) dy = \frac{34}{105} \approx 0.32381$$

Integrate[y^2 (6y^5 - 10y^4 + 4y^3 - 2y + 2), {y, 0, 1}]

Extended Keyboard
Upload
Examples
Random


Definite integral:
More digits
Step-by-step solution

$$\int_0^1 y^2 (6y^5 - 10y^4 + 4y^3 - 2y + 2) dy = \frac{13}{84} \approx 0.15476$$


Hence,  $E(Y) = 0.32381$ ,  $Var(Y) = 0.15476 - 0.32381^2 = 0.049907084$ .

6. ...

Integrate[z (6z^5), {z, 0, 1}]

 Extended Keyboard

 Upload

 Examples

 Random

Definite integral:

[More digits](#)


☒ Step-by-step solution

$$\int_0^1 z (6z^5) dz = \frac{6}{7} \approx 0.85714$$

Integrate[z^2 (6z^5), {z, 0, 1}]

 Extended Keyboard

 Upload

 Examples

 Random

Definite integral:

☒ Step-by-step solution

$$\int_0^1 z^2 (6z^5) dz = \frac{3}{4} = 0.75$$

Hence,  $E(Z) = 0.85714$ ,  $Var(Z) = 0.75 - 0.85714^2 = 0.01531102$ .

7. Dividing  $f(x, y, z)$  by the marginal pdf of  $X$  and  $Z$  yields the conditional density of  $Y$ . We can then derive the expectation by definition.


$$\begin{aligned} f_{Y|X,Z}(y|x,z) &= \frac{f(x,y,z)}{\text{margin}(x) \cdot \text{margin}(z)} \\ &= \frac{8(x^3 + y^3 + z^3)}{(-12x^5 + 10x^4 + 2x) \cdot (6z^5)} \\ E(Y|X=x, Z=z) &= \int_0^1 y \cdot f_{Y|X,Z}(y|x,z) dy \\ &= \int_0^1 \frac{8(x^3 + y^3 + z^3) \cdot y}{(-12x^5 + 10x^4 + 2x) \cdot 6z^5} dy \end{aligned}$$




Integrate[8y (x^3 + y^3 + z^3) / (6z^5 (-12x^5 + 10x^4 + 2x)), {y, 0, 1}]

 Extended Keyboard

 Upload

 Examples

 Random

Definite integral:

 Step-by-step solution

$$\int_0^1 \frac{8y(x^3 + y^3 + z^3)}{6z^5(-12x^5 + 10x^4 + 2x)} dy = \frac{5x^3 + 5z^3 + 2}{15(-6x^5 + 5x^4 + x)z^5}$$

### 3 Bayesian decision theory

1.  $\omega_1$  has 6 data points,  $\omega_2$  has 8 data points, so the prior probability of each class is:

$$P(\omega_1) = 6/14 = 3/7 = 0.428571429, P(\omega_2) = 8/14 = 4/7 = 0.571428571$$

2. By definition, we can calculate that:

```
1 import numpy as np
2
3 w1 = np.array([[0, 0], [1, 1], [2, 2], [3, 2], [4, 2], [5, 3]])
4 w2 = np.array([[6, 9], [8, 10], [9, 10], [9, 11], [10, 10], [11, 12], [12, 12], [12,
5     13]])
6
7 def mean(w):
8     return np.array([np.sum(w[:,0]) / w.shape[0], np.sum(w[:,1]) / w.shape[0]])
9
10 def sigma(w):
11     dof = w.shape[0] - 1 # degree of freedom
12     deviation = (w - mean(w))
13     return 1 / dof * np.dot(deviation.T, deviation)
14
15 if __name__ == '__main__':
16     for idx, w in enumerate([w1, w2]):
17         with np.printoptions(precision=3, suppress=True):
18             print(f'mu_{idx + 1} = {mean(w)}')
19             print(f'sigma_{idx + 1} = {sigma(w)}')
20
21     '''results'''
22     # mu_1 = [2.5, 1.667]
23     # sigma_1 = [3.5 1.8]
24     # [1.8 1.067]
25     #
26     # mu_2 = [9.625, 10.875]
27     # sigma_2 = [4.268 2.518]
28     # [2.518 1.839]
```

test.py

However, if the data size is large, the functions are very computationally expensive, we should use vectorized functions to boost performance.

```
1 mu = np.mean(w, axis=0)
2 sigma = np.cov(w.T)
```

3. Since the covariance matrices for the two classes are neither identical nor  $\sigma^2 I$ , this problem falls into the case where sigma is arbitrary. In particular, the prior probabilities are not uniform, so our decision is based on both priors and likelihood. As we did in the exercise, the discriminant function can be formulated as:

$$g_i(x) = \ln p(x|\omega_i) + \ln P(\omega_i)$$

$$g_i(x) = -\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1}(x - \mu_i) - \frac{1}{2} \ln |\Sigma_i| + \ln P(\omega_i)$$

In Python, we then compute the inverse and determinant of the covariance matrix:

```
1 import numpy as np
2
3 w1 = np.array([[0, 0], [1, 1], [2, 2], [3, 2], [4, 2], [5, 3]])
4 w2 = np.array([[6, 9], [8, 10], [9, 10], [9, 11], [10, 10], [11, 12], [12, 12], [12,
5     13]])
6 if __name__ == '__main__':
7     for w in [w1, w2]:
8         sigma = np.cov(w.T)
9         sigma_inv = np.linalg.inv(sigma)
10        print('inverse matrix = ', np.linalg.inv(sigma))
11        print('determinant = ', np.log(np.linalg.det(sigma)))
12
13        '''result'''
14        # inverse matrix = [2.16216216 - 3.64864865]
15        #                   [-3.64864865  7.09459459]
16        # determinant = -0.7065702008920836
17        # inverse matrix = [1.21790541 - 1.66722973]
18        #                   [-1.66722973  2.82601351]
19        # determinant = 0.4122447950935426
```

test1.py

Therefore, we have:

$$\{x-2.5, y-1.667\} \cdot \{[2.162, -3.648], [-3.648, 7.094]\} \cdot \{[x-2.5], [y-1.667]\} + 0.707/2 + \ln(3/7)$$

Extended Keyboard

Upload

Examples

Random

Input:

$$\{x-2.5, y-1.667\} \cdot \begin{pmatrix} 2.162 & -3.648 \\ -3.648 & 7.094 \end{pmatrix} \cdot \{x-2.5, y-1.667\} + \frac{0.707}{2} + \log\left(\frac{3}{7}\right)$$

log(x) is the natural logarithm

Result:

$$(x-2.5)(2.162(x-2.5)-3.648(y-1.667)) + (y-1.667)(7.094(y-1.667)-3.648(x-2.5)) - 0.493798$$

Expanded form:

$$2.162x^2 - 7.296xy + 1.35243x + 7.094y^2 - 5.4114y + 2.32606$$

$$g_1(x)$$



$$\{x-9.625, y-10.875\} \cdot \{[1.217, -1.667], [-1.667, 2.826]\} \cdot \{[x-9.625], [y-10.875]\} - 0.412/2 + \ln(4/7)$$

Extended Keyboard

Upload

Examples

Random

Input interpretation:

$$\{x - 9.625, y - 10.875\} \cdot \begin{pmatrix} 1.217 & -1.667 \\ -1.667 & 2.826 \end{pmatrix} \cdot \{x - 9.625, y - 10.875\} - \frac{0.412}{2} + \log\left(\frac{4}{7}\right)$$

$\log(x)$  is the natural logarithm

Result:

$$(x - 9.625)(1.217(x - 9.625) - 1.667(y - 10.875)) + (y - 10.875)(2.826(y - 10.875) - 1.667(x - 9.625)) - 0.765616$$

Expanded form:

$$1.217x^2 - 3.334xy + 12.83x + 2.826y^2 - 29.3758y + 97.2206$$

$$g_2(x)$$

The decision boundary is derived by setting:  $g_1(x) = g_2(x)$ , which is equivalent to the ellipse function computed by Mathematica:

$$0.945x^2 - 3.96xy - 11.478x + 4.268y^2 + 23.96y - 94.89 = 0 \Rightarrow$$

$$y = \frac{990x - 5990 \pm \sqrt{-28215x^2 + 386826x + 137127730}}{2134}$$

Finally, we visualize the scatter plot of the two classes and draw the decision boundary in 2D.

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 w1 = np.array([[0, 0], [1, 1], [2, 2], [3, 2], [4, 2], [5, 3]])
5 w2 = np.array([[6, 9], [8, 10], [9, 10], [9, 11], [10, 10], [11, 12], [12, 12], [12,
6     13]])
7 if __name__ == '__main__':
8     scatter1 = plt.scatter(w1[:, 0], w1[:, 1], c='red', s=1)
9     scatter2 = plt.scatter(w2[:, 0], w2[:, 1], c='blue', s=1)
10
11     x = np.linspace(-80, 80, 500)
12     y1 = (990 * x - 5990 + np.sqrt(-28215 * (x ** 2) + 386826 * x + 137127730)) / 2134
13     y2 = (990 * x - 5990 - np.sqrt(-28215 * (x ** 2) + 386826 * x + 137127730)) / 2134
14     plt.plot(x, y1, color='green', lw=1)
15     boundary, = plt.plot(x, y2, color='green', lw=1)
16
17     plt.title('decision boundary of w1, w2')
18     plt.xlabel('x')
19     plt.ylabel('y')
20     plt.grid(alpha=.6, linestyle='--')
21     plt.legend([scatter1, scatter2, boundary],
22               ['w1 class', 'w2 class', 'Bayes decision boundary'],
23               loc='lower right')
24     plt.show()
```

test2.py



4. If the penalties for misclassification are different for the two classes, it will definitely affect the decision boundary. In the case of 2-category classification, the Bayes decision rule is based on the likelihood ratio, so we would decide  $\omega_1$  if

$$\frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} > \frac{\lambda_{12} - \lambda_{22} P(\omega_2)}{\lambda_{21} - \lambda_{11} P(\omega_2)}$$

The left hand side is the likelihood ratio, the right hand side is the threshold value, which can vary drastically depending on how we define the loss function  $\lambda()$ . For instance, if our loss function penalizes miscategorizing  $\omega_2$  as  $\omega_1$  more than the converse ( $\lambda_{12} > \lambda_{21}$ ), we end up with a larger threshold, and hence the decision regions change as well, as shown in the figure below.

