

COMP 532

Machine Learning and Bioinspired Optimisation

Lecture 3



UNIVERSITY OF
LIVERPOOL

Meng Fang

University of Liverpool

Overview

- Mathematical Preliminaries
- Python Basics

Overview

- Mathematical Preliminaries
- Python Basics

Probability basics

- A random variable X represents outcomes or states of the world
- We will write $p(x)$ to mean $\text{Probability}(X = x)$
- Sample space: The space of all possible outcomes (may be discrete, continuous, or mixed)
- $p(x)$ is the probability mass (density) function
 - Assigns a number to each point in sample space
 - Non-negative, sums (integrates) to 1
 - Intuitively: how often does x occur, how much do we believe in x

Probability basics

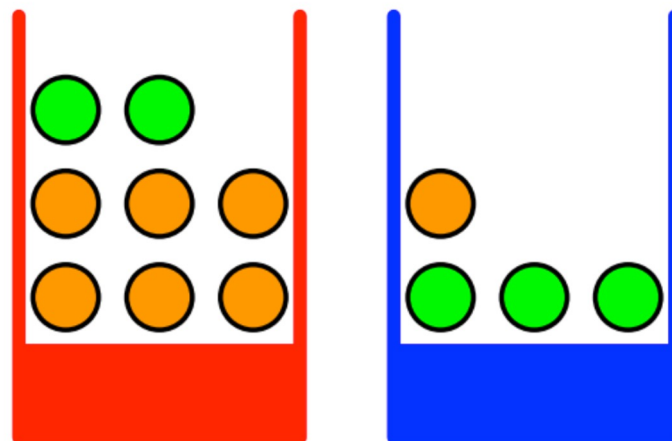
- Random Experiment – Experiment that results in different outcomes despite being in similar conditions
 - Tossing of a coin, throwing of a dice, rainfall amount
- Sample Space: Set of all possible outcomes of a random experiment
 - Tossing of a coin once: $S = \{H, T\}$ (head and tail)
 - Tossing coin twice: $S = \{HH, HT, TH, TT\}$

Probability basics

- Joint Probability
- Sum Rule
- Conditional Probability
- Product Rule
- Bayes' Rule

A Simple Example

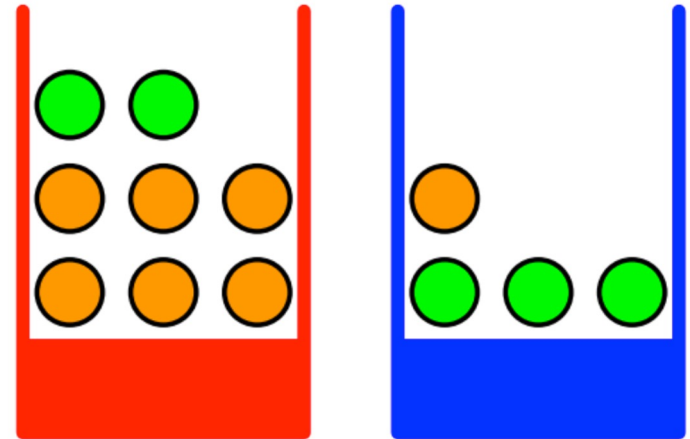
- Two Baskets – Red & Blue
- Each basket has some fruits – oranges and apples (green)
- All fruits in a basket are equally available
- Red Basket - 6 oranges, 2 apples
- Blue Basket – 1 orange, 3 apples
- Let $P(B=r) = 0.4$ and $P(B=b) = 0.6$
- Random variables:
 - $B: \{b, r\}$; $F: \{o, a\}$



- What is the probability of picking an orange?
- What is the probability that I pick red basket given that the fruit I pick was an orange?

A Simple Example

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$



	$F = o$	$F = a$	
$B = r$	$6/8 \times 40 = 30$	$2/8 \times 40 = 10$	40
$B = b$	$1/4 \times 60 = 15$	$3/4 \times 60 = 45$	60
	45	55	

Joint Probability

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$

	F = o	F = a	
B = r	30	10	40
B = b	15	45	60
	45	55	

- Joint Probability: The probability (that X will take a value x_i and Y will take a value y_j) $P(X = x_i, Y = y_j)$

- $P(B=r, F=o) = 30/100 = 0.3$

y_j

		x_i		
		n_{ij}		

c_i

Generalize

- Let the number of trials that $X = x_i$ and $Y = y_j$ be n_{ij} .

Then, $P(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$, where N is the total number of trials

- $P(B=b, F=a) = ?$

Sum Rule

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$

	F = o	F = a	
B = r	30	10	40
B = b	15	45	60
	45	55	

- What is the probability that I pick oranges?
 - $P(F=o) = 45/100 = 0.45$

		x_i		
		n_{ij}		

y_j

c_i

Generalize

- Then, $P(X = x_i) = \frac{c_i}{N}$ (marginal probability)
- However, $c_i = \sum_j n_{ij}$
- Then, $P(X = x_i) = \sum_j \frac{n_{ij}}{N}$

Sum Rule

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$

	F = o	F = a	
B = r	30	10	40
B = b	15	45	60
	45	55	

- What is the probability that I pick oranges?
 - $P(F=o) = 45/100 = 0.45$

		x_i		
		n_{ij}		

y_j

c_i

Generalize

- Then, $P(X = x_i) = \frac{c_i}{N}$ (marginal probability)
- However, $c_i = \sum_j n_{ij}$
- Then, $P(X = x_i) = \sum_j \frac{n_{ij}}{N}$
- We have $P(X = x_i, Y = y_i) = \frac{n_{ij}}{N}$
- Then $P(X = x_i) = \sum_j P(X = x_i, Y = y_i) \leftarrow$ Sum rule of Probability

Conditional Probability

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$

	F = o	F = a	
B = r	30	10	40
B = b	15	45	60
	45	55	

- What is the probability that the basket I picked is red given that the fruit was orange: $P(B = r|f=o) = 30/45$ x_i

A diagram illustrating a matrix structure. It consists of a 3x5 grid of cells. The columns are labeled x_i at the top, and the rows are labeled y_j on the left. The element in the second row and third column is labeled n_{ij} . A blue bracket under the third and fourth columns is labeled c_i .

- What is the probability that Y will take the value y_j given that X will take the value x_i
- Then $P(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i}$

Product Rule

- Assume $N = 100$ trials
- $P(B=r) = 0.4$ and $P(B=b) = 0.6$

	F = o	F = a	
B = r	30	10	40
B = b	15	45	60
	45	55	

- Conditional Probability

- $P(Y = y_i | X = x_i) = \frac{n_{ij}}{c_i}$

- Joint Probability

- $P(X = x_i, Y = y_i) = \frac{n_{ij}}{N}$

- We can write $\frac{n_{ij}}{N} = \frac{n_{ij}}{c_i} \frac{c_i}{N}$ (We also have $P(X = x_i) = \frac{c_i}{N}$)

- Thus $P(X = x_i, Y = y_i) = P(Y = y_i | X = x_i) P(X = x_i)$

x_i

		n_{ij}		

y_j

c_i

Rules Simplified

- Sum rule $P(X = x_i) = \sum_j P(X = x_i, Y = y_j)$
- Produce rule $P(X = x_i, Y = y_i) = P(Y = y_i | X = x_i) P(X = x_i)$

Simply,

- Sum rule $P(X) = \sum_y P(X, Y)$
- Produce rule $P(X, Y) = P(Y|X) P(X)$

Bayes' Rule

- Produce Rule $P(A,B)=P(A|B)P(B)$
- Similarly, $P(B,A)=P(B|A)P(A)$
- Since $P(A,B)= P(A|B)P(B)$
- Then we have $P(A|B)P(B) = P(B|A)P(A)$

$$P(A|B) = P(A) \frac{P(B|A)}{P(B)}$$

Posterior Prior Likelihood
Marginal

Python Basics

- Python is an interpreted, high-level and general- purpose programming language.
 - Created by Guido van Rossum and first released in 1991
- Has a lot of : <https://www.python.org>
 - Image Processing
 - Computer Vision
 - Machine Learning
 - Deep Learning
 - Optimization
 - Signal Processing
 - ...

Why Python

- Open source
 - Ease of coding, “Code as plain English” is Python’s primary goal.
 - Most commonly used, a bigger community. There are a lot of blog posts and online resources regarding Python + OpenCV
 - Vast libraries for machine learning, computer vision
 - ...
-
- Python vs. C++ vs. Java vs. Matlab

Tools to use Python

- **Anaconda**
 - Anaconda refers to an open source Python distribution, which includes more than 180 scientific packages and their dependencies, e.g., conda, Python, Jupyter Notebook.
- **VSCode/Pycharm**
 - An Integrated Development Environment (IDE) for Python. It has a set of tools that can help users improve their efficiency when developing in Python language.
Such as debugging, syntax highlighting, project management, code jump, smart prompt, auto completion, etc.
- **Jupyter Notebook**
 - Jupyter Notebook is a web-based application for interactive computing. It can be applied to the whole process of development: code writing, running the code and displaying the results.
 - Install the package directly in Jupyter Notebook, such as: `pip install pytorch`

Tools to use Python

- Google Colab

- Colaboratory is a free Jupyter Notebook environment that requires no setup and runs on the Cloud (write, run and share code)
- No configuration required
- Free use of GPU, TPU
- Easy sharing

Install the package directly in Jupyter Notebook, such as: `pip install tensorflow`

- Online platform: <https://colab.research.google.com/notebooks/intro.ipynb>
 - Similar to Jupyter Notebook

How to start

Data structures in Python

- Basic

- int
- float
- bool

- Ordered

- list: [2, 3, 3]
- array: `np.array([2, 3, 3])` # using numpy
- tuple: (1,3,6,10)
- string: 'hello'

- Unordered

- set: `set('abracadabra')`
- dict: `{'jack': 4098, 'sape': 4139}`

How to start

Useful tricks

- Array

- Vector is stored as 1D array `[1 2 3 4 5 6]`
- Matrix is 2D array `[[1 2 3]
[4 5 6]]`

- The index is from 0

- For example, let `A = [1,2,3,4,5,6]` then `A[0] = 1`

- Array in Numpy

```
import numpy as np
```

```
a1 = np.array([3, 4, 5, 3])
```

- Matrix in Numpy

```
import numpy as np
```

```
m1 = np.random.rand(2,2)
```

How to start

Function in Python

- Creating a function: In Python a function is defined using the def keyword.

```
def my_function():  
    print("Hello from a function")
```

- Calling a function

```
def my_function():  
    print("Hello from a function")  
  
my_function()
```

- Arguments

```
def my_function(fname):  
    print(" My name is " + fname)  
  
my_function("Shan")
```

How to start

Syntax in Python

- if...elif...else

```
a = 200
b = 33
if b > a:
    print("b is greater than a")
elif a == b:
    print("a and b are equal")
else:
    print("a is greater than b")
```

- for

```
fruits = ["apple", "banana", "cherry"]
for x in fruits:
    print(x)
```

- while

```
i=1
while i < 6:
    print(i) i += 1
```

How to start

Debugging in Python

- Debugging with pdb: easy, powerful, no extra IDE needed

<https://www.youtube.com/watch?v=bHx8A8tbj2c>

- Debugging with PyCharm: easy, but not free

More you can explore

Check out more in w3schools, and practice!

<https://www.w3schools.com/python/>

Summary

- Probability basics
- Introduction to Python