



Seoul National University
College of Engineering
Department of Naval Architecture and Ocean Engineering
1, Gwanak-ro, Gwanak-gu, Seoul 151-744, Korea

Fall 2020

머신러닝

PA # 1

Instructor name	김태완 교수님
Student name	이용준
Department	조선해양공학과
Student ID	2015-19595
Submission date	2020.09.27

Contents

1. Problem Definition	3
1.1 Problem 1	3
1.2 Problem 2	3
2. Problem Analysis and Design	3
2.1 Analysis	3
2.2 Data Flow Diagram	4
3. Code Explanation	5
3.1 Class	5
3.2 Function	5
4. Conclusion	7
4.1 Result	7
4.2 Conclusion	9

1. Problem Definition

1.1 Problem 1

- 10개의 삼각형 좌표에 대한 정보를 `input.txt`로 만든다.
- 이를 읽어와 삼각형 넓이를 계산한다.
- 넓이에 대한 평균, 표준편차를 계산하고 결과를 `output.txt`로 출력한다.
- 넓이의 도수분포를 계산하여 `matplotlib`을 이용하여 `plot`한다.

1.2 Problem 2

- 문제 1의 삼각형에 대하여 `affine` 변환을 시킨 새로운 삼각형들의 좌표를 계산한다.
- 변환된 좌표를 파일로 출력한다.
- 변환 전과 후의 삼각형 그림을 `matplotlib`을 이용하여 `plot`한다.

2. Problem Analysis and Design

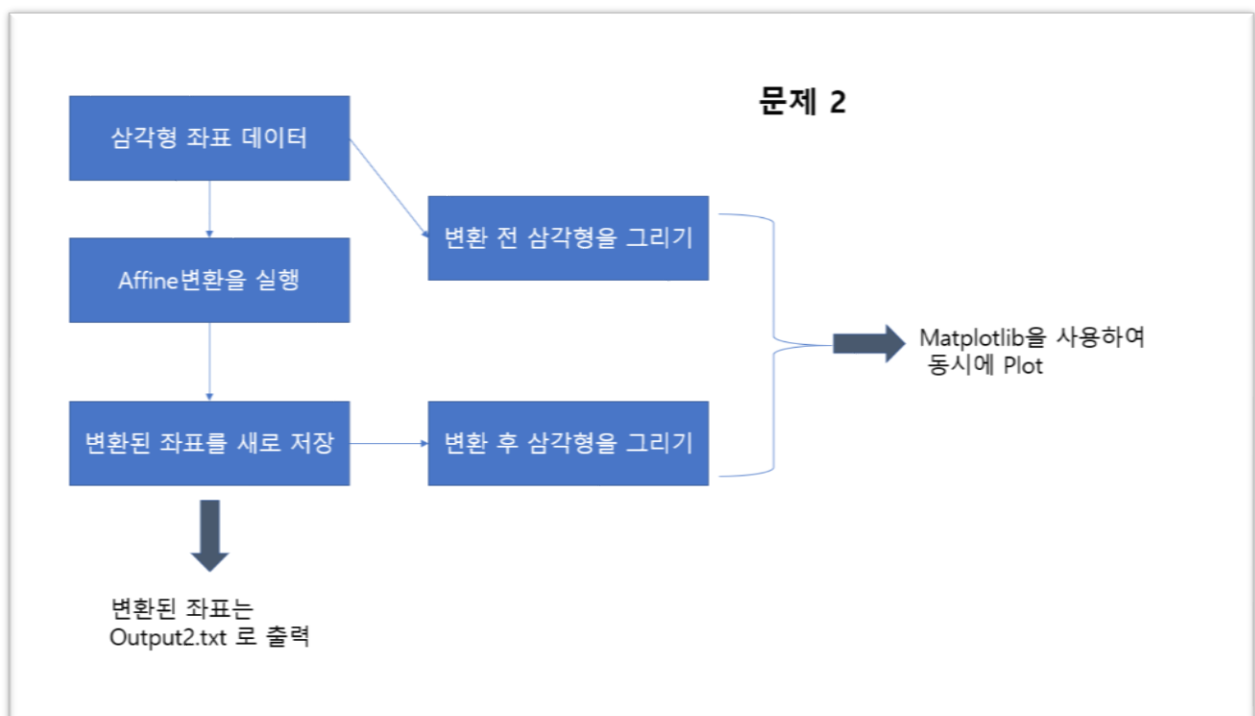
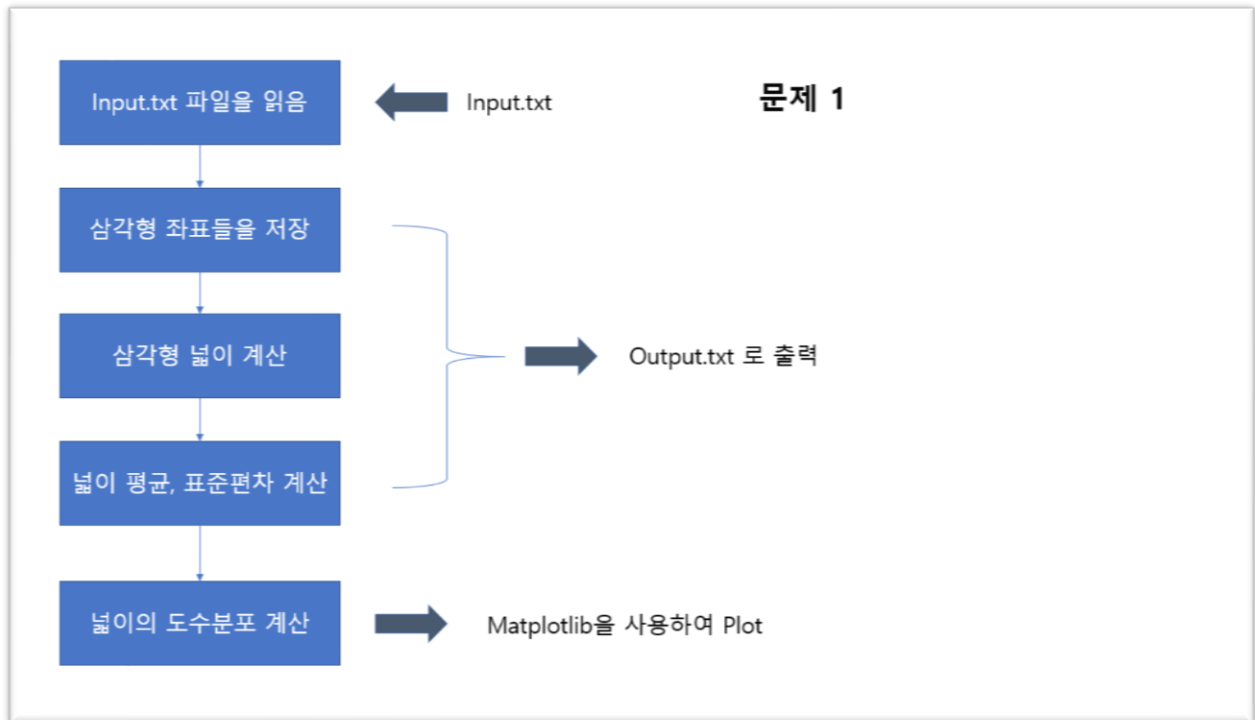
2.1 Analysis

- 이번 과제를 수행하기 위해선 외부의 텍스트 파일을 프로그램 상에서 읽고 쓸 줄 알아야 한다. 파이썬을 이용하면 비교적 쉽게 텍스트 파일을 읽고 쓸 수 있게 된다. 알맞은 경로에 `input.txt` 파일을 만들고 10개의 삼각형에 대한 좌표를 작성한다. 한 개의 삼각형을 구성하는 3개의 점을 구성하는 6개의 좌표를 순서대로 한 줄에 작성한다. 이를 프로그램 상에서 읽어와 좌표 정보를 배열 형태로 저장한다.
- 간단한 공식을 이용하면, 좌표들을 이용하여 삼각형 넓이를 계산할 수 있다. 또한 삼각형 넓이의 표준편차는 편차 제곱의 평균에 제곱근을 씌워 구할 수 있다. 이때 편차는 평균과의 편차를 의미하며, 나누는 수는 문제의 조건에 맞게 $N-1$ 로 나누어 준다.
- `Affine` 변형은 회전변환과 평행이동이 합쳐진 변환이다. 이는 행렬 곱 연산과 행렬 합 연산을 이용하여 어렵지 않게 변환된 좌표를 계산할 수 있다. `Affine` 변환에 사용되는 행렬의 성분들은 임의로 정한다.
- 도수분포를 그리기 위하여 `matplotlib`에 있는 `hist()` 함수를 이용한다. 이때 임의로 x축에 해당하는 값의 범위랑 간격을 조정할 수 있다.
- 문제 2를 수행하기 위해서는 한번에 여러 개의 그림을 동시에 그릴 수 있어야 한다. `Matplotlib`의 `GridSpec`을 사용하면 임의로 정한 행 수, 열 수만큼 구역이 나뉘지게 되어, 그

구역에 subplot을 할 수 있게 된다.

2.2 Data Flow Diagram

각 문제에 대한 DFD이다.



3. Code Explanation

3.1 Class

프로젝트에 사용된 클래스에 관한 설명이다.

<문제 1>

Class Name	Explanation
c_read_file	Input 파일을 읽는다.
c_triangle_2D	삼각형 좌표들을 저장한다.
c_calculator	삼각형의 넓이, 평균, 표준편차를 계산한다.
c_draw_histogram	넓이의 도수분포를 계산하여 그린다.
c_write_file	삼각형 넓이, 평균, 표준편차를 output파일에 출력한다.

<문제 2>

Class Name	Explanation
c_affine_transform	Affine 변환을 실행한다.
c_matrix	2x2 Matrix를 초기화한다.
c_vector	2차원 벡터를 초기화한다.
c_draw_triangle	좌표 정보를 이용하여 삼각형들을 그려낸다.

3.2 Function

각 클래스별 함수에 대한 설명이다.

c_read_file		
read_input_data(file_name)		
Parameter	Type	Explanation
file_name	string	파일의 이름을 문자열로 저장

c_triangle_2D		
init_triangle_info(all_lines, number_of_lines)		
Parameter	Type	Explanation
all_lines	array	텍스트 파일의 한 줄씩을 각각 저장한 배열
number_of_lines	int	텍스트 파일의 줄 수

c_calculator		
cal_area_of_triangle(a)		
Parameter	Type	Explanation
a	array	하나의 삼각형에 대한 좌표 성분 6개를 배열로 저장
cal_average(list)		
Parameter	Type	Explanation
list	array	삼각형들의 넓이를 배열 형태로 저장
cal_standard_deviation(list)		
Parameter	Type	Explanation
list	array	삼각형들의 넓이를 배열 형태로 저장

c_draw_histogram		
draw_histogram(area_of_triangle)		
Parameter	Type	Explanation
area_of_triangle	array	삼각형들의 넓이를 배열 형태로 저장

c_write_file		
write_output_data(area_of_triangle, mean_of_area, standard_deviation)		
Parameter	Type	Explanation
area_of_triangle	array	삼각형들의 넓이를 배열 형태로 저장
mean_of_area	Float	삼각형 넓이의 평균
standard_deviation	float	넓이의 표준편차
write_vector_of_affine_transformed_triangle(transformed_triangle)		
Parameter	Type	Explanation
transformed_triangle	array	Affine 변환된 삼각형들의 좌표가 배열형태로 저장

c_affine_transform		
calculate_affine_transform(x,y, matrix , vector)		
Parameter	Type	Explanation
x	float	Affine 변환에 대입시킬 좌표의 x 성분
y	float	Affine 변환에 대입시킬 좌표의 y 성분
matrix	array	회전 변환에 해당하는 행렬을 배열로 저장
vector	array	평행 이동에 해당하는 벡터를 배열로 저장

c_matrix		
set_matrix(theta)		
Parameter	Type	Explanation
theta	float	삼각형을 회전시킬 각도 theta

c_vector		
set_vector(a, b)		
Parameter	Type	Explanation
a	Float	2차원 벡터의 x 성분
b	float	2차원 벡터의 y 성분

c_draw_triangle		
draw_triangles(triangle, location)		
Parameter	Type	Explanation
triangle	array	삼각형들의 좌표 성분들을 저장한 배열들을 담고 있는 배열
location	array	Sub plot을 그릴 위치 정보를 배열로 저장

4. Conclusion

4.1 Result

문제 1에 대한 결과로, 삼각형들의 넓이와 평균, 표준편차, 도수분포 정보를 담고 있는 output 파일과 도수분포를 그린 히스토그램 그림 파일이 나온다. 다음은 그 결과이다.

```

output_data - Windows 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
output
Triangle[0] = 6.000000
Triangle[1] = 4.000000
Triangle[2] = 1.000000
Triangle[3] = 2.000000
Triangle[4] = 7.500000
Triangle[5] = 2.500000
Triangle[6] = 1.500000
Triangle[7] = 7.500000
Triangle[8] = 1.000000
Triangle[9] = 2.000000
Average of Area : 3.5
Standard Deviation of Area : 2.592724864350674

Frequency Distribution of Area
0~2 : 3
2~4 : 3
4~6 : 1
6~8 : 3
8~10 : 0
  
```

그림 1. 삼각형 넓이와 평균, 표준편차, 도수분포

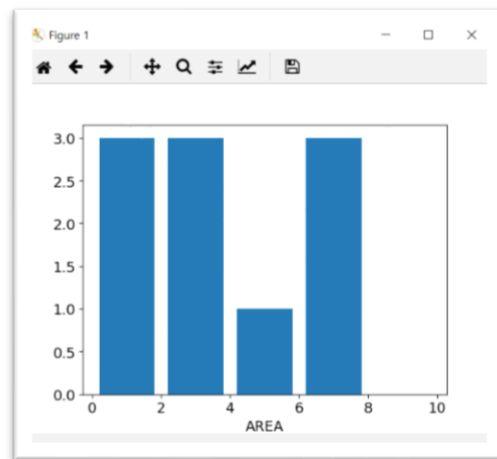


그림 2. 넓이에 대한 도수분포

그리고 문제2에 대한 결과로, affine 변환된 삼각형들의 새로운 좌표 정보를 담고 있는 output 파일 하나와, 변환 전 후의 삼각형 그림이 있다. 다음은 그 결과이다.

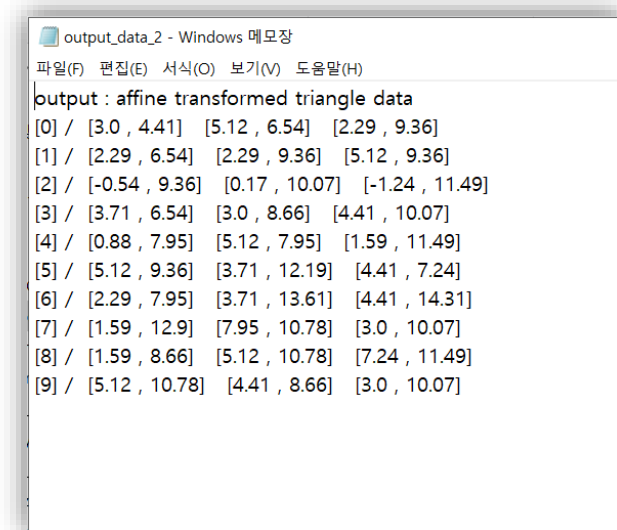


그림 3. Affine 변환된 새로운 삼각형들의 좌표

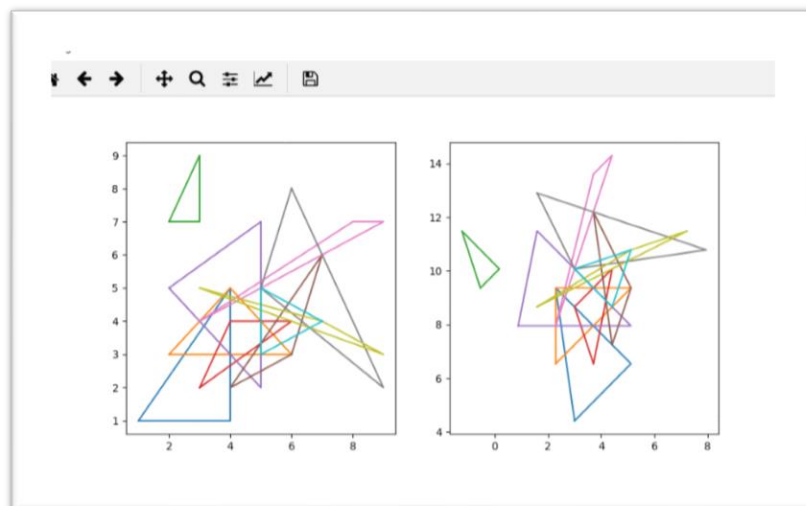


그림 4. Affine변환 전후의 삼각형 그림

4.2 Conclusion

- 텍스트 파일을 활용하여 더욱 쉽게 입출력 결과를 확인할 수 있었다.
- 여러 클래스들을 활용하여 각 함수를 기능별로 간편하게 분류해서 사용할 수 있었다.
- 기존 함수를 사용하지 않고 직접 계산으로 표준편차를 구해볼 수 있었다.
- Matplotlib을 활용하여 도수분포를 히스토그램으로 그림으로써 분포도를 한눈에 확인할 수 있었다.
- Matplotlib을 활용하여 여러 개의 그래프를 한번에 그림으로써 affine transformation을 더욱 잘 확인할 수 있었다.
- PA#1을 수행함으로써 파이선에 대한 기초 문법과 함수 사용법, 그림 그리는 방법 등을 배울 수 있었다.