B-IT Pattern Recognition

Project-I Report

Students: *Mohammadali Ghasemi, Chu-I Chao, Neng Qian, Vanik Karslian, Can Yuce*

Bonn-Aachen International Center for
Information Technology(B-IT)

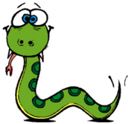
November 29, 2017

Outline

Generalities

Programming Language, Libraries:

- Implementation of the solution is in Python.
- Using NumPy, SciPy, Matplotlib modules.
- Anaconda data science platform[1] (not obligatory).



[1] https://www.anaconda.com/

Task 1.1: Plotting 2D data

- **Task:** Plotting the data without the outliers,
- **Why?**
  - To deal with missing data,
  - To deal with outliers.
- **How?**

A Python Code Sample:

```python
# Task 1.1
outlierInd = np.where( X[:,0] != -1 )
X, y = X[outlierInd], y[outlierInd]
```

Task 1.1: Plotting 2D data

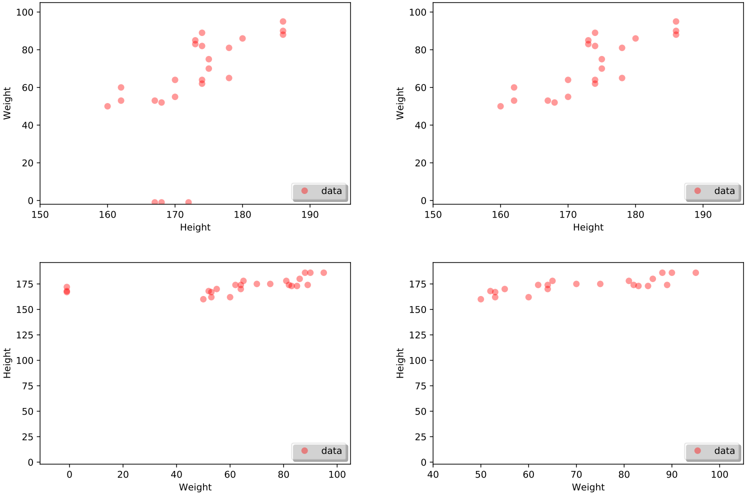


Figure 1: Removing outliers.

Task 1.1: Plotting 2D data

```
# Task 1.1
axs.set_aspect('equal')
```

- same scaling from data to plot units for x and y,

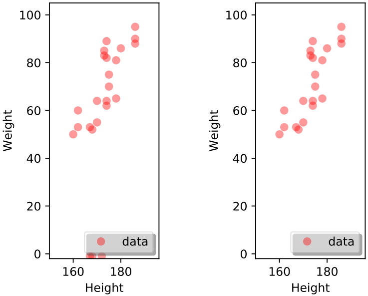


Figure 2: Plots with equal aspect ratio.

Task 1.2: Fitting a Normal Distribution on 1D Data

- **Task:**
  - Calculate the mean and standard deviation of body sizes,
  - Plot the data and normal distribution.
- **Why?**
  - To be sure that body sizes are distributed accordingly to normal distribution.
- **How?**
  - Using ML estimation:

$$\text{Normal Distribution: } f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Standard Deviation: } \sigma_{MLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\text{Mean: } \mu_{MLE} = \frac{1}{n}\sum_{i=1} ns_i$$

Task 1.2: Fitting a Normal Distribution on 1D Data

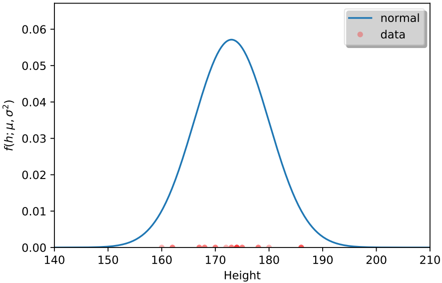


Figure 3: $\mu_h = 173.0, \sigma_h = 6.98$

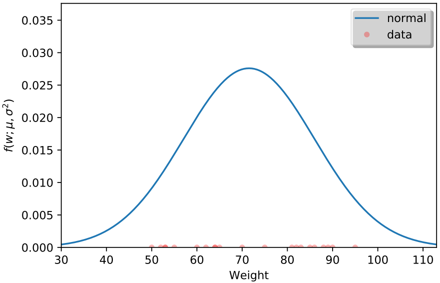


Figure 4: $(\mu_w = 71.52, \sigma_w = 14.46)$

Task 1.3: Fitting Weibull Distribution to 1D Data

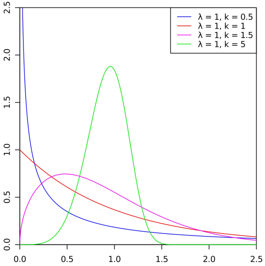


Figure 5: Weibull Distribution.

PDF of Weibull distribution:

$$f(x|\kappa,\alpha) = \frac{\kappa}{\alpha}\left(\frac{x}{\alpha}\right)^{\kappa-1} e^{-\left(\frac{x}{\alpha}\right)^{\kappa}}, \text{ if } x \geq 0 \text{ otherwise } 0.$$

where κ is the shape parameter and

α is the scale parameter.

Task 1.3: Fitting Weibull Distribution to 1D Data

- **Task:**
  - Estimate the parameters of Weibull distribution,
  - Plot the data and Weibull distribution..
- **Why?**
  - To have a good prediction if the Weibull distribution can be closely fitted to the observed data.
- **How?**
  - Maximum Likelihood Estimation(MLE): Estimate κ, α which makes the log-likelihood as large as possible.
  - Use the function scipy.integrate.odeint.

Task 1.3: Fitting Weibull Distribution to 1D Data

- **How?**
  - Maximum Likelihood Estimation(MLE): Estimate κ, α which makes the log-likelihood as large as possible.

    PDF of Weibull distribution:

    $$f(x|\kappa, \alpha) = \frac{\kappa}{\alpha} \left(\frac{x}{\alpha} \right)^{\kappa-1} e^{-\left(\frac{x}{\alpha}\right)^{\kappa}},$$

    Log-likelihood is:

    $$L(\alpha, \kappa|D) = N\left(\log \kappa) - \kappa \log(\alpha)\right) + (\kappa - 1) \sum_i \log d_i - \sum_i \left(d_i/\alpha\right)^{\kappa}$$

  - Use Newton Raphson Method:

    $$\begin{bmatrix} \kappa^{\text{new}} \\ \alpha^{\text{new}} \end{bmatrix} = \begin{bmatrix} \kappa \\ \alpha \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 L}{\partial \kappa^2} & \frac{\partial^2 L}{\partial \kappa \partial \alpha} \\ \frac{\partial^2 L}{\partial \kappa \partial \alpha} & \frac{\partial^2 L}{\partial \alpha^2} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial L}{\partial \kappa} \\ -\frac{\partial L}{\partial \alpha} \end{bmatrix}$$

  - Run for 20 iterations.

Task 1.3: Fitting Weibull Distribution to 1D Data

- Plot the data histogram with the scaled version of the fitted distribution.

```
def model(y, t):
        K, a = y[0], y[1]
        d_K = N / K - N * math.log(a) + sum_log_di - np.sum
            ( ((hist / a) ** K) *  np.log(hist / a))
        d_a = (K / a) * (np.sum((hist / a) ** K) - N)
        return [d_K, d_a]

# initial condition and  time points
K, a = 1., 1.
t = np.linspace(0,100,1000)
y = odeint(model,[K,a],t)
```

Task 1.3: Fitting Weibull Distribution to 1D Data

- Plot the data histogram with the scaled version of the fitted distribution.

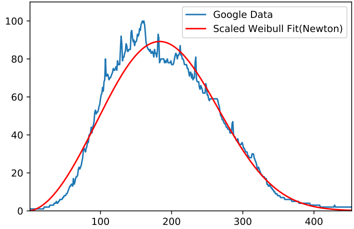


Figure 6: Data fitting with Newton Raphson Method(left) and scipy.integrate.odeint(right).