Homework #2

Due Monday, Feb 14, 2023, 11:59pm

Problem 2.1.

a) Read in the 2 column data from the file 'bfit.csv' into an np.array()
b) Print a list of the matrix, and plot the points as dots in magenta color. The x-axis is the first column, the y-axis is the second column.
c) Calculate the bounding box, and plot a red rectangle around it
d) Draw the diagonal from (xmin,ymin)-(xmax,ymax) in blue
e) Overplot the points below the diagonal in green
f) Calculate the center of mass of the points in the two halves separated by the diagonal. Plot these values shown with an asterisk on the same figure.

Problem 2.2.

a) Create a uniform array x with 101 elements between 0 and 2*pi
b) Create an array containing y = sin(3*x)
c) Create a plot y vs x
d) Create another array, z = y*y
e) Plot z vs x
f) Calculate the average of both y and z over this interval
g) How do the results of (f) change if we use 10000 points?

Problem 2.3.

a) Create a 21x21 grid of x and y values over a square [-1,1]x[-1,1]
b) Write a function that is a Gaussian,

$$\text{fgauss}(x, y, s) = \frac{1}{\sqrt{2\pi s^2}} \exp\left(-\frac{x^2 + y^2}{2s^2}\right)$$

c) Create a contour plot of z as a function of x,y for the values of s=1,2,3

Problem 2.4.

Consider the data in the files `a100.csv, b100.csv, c100.csv` and `d100.csv`.

a. Determine the underlying probability distributions (and its parameters) of each data set, by creating a histogram and over-plotting with the most similar probability distribution, until the agreement is acceptable. Create a label with the name of the distribution, and its parameter values on the plot. Do not use a fitting function but determine the parameters by changing them manually until there is a good visual match. The goal of this exercise is to develop an intuition on how the shapes of the different distributions change as a function of the parameters.

b. Create a new series from each data set through the formula

$$y_p = \sum_{i=0}^{K-1} x_{p+i}$$

i.e. each new number is the sum of K adjacent elements of the original series (so called moving average). Determine the probability distribution and its parameter for each sequence for K=5, 20 and 80. Calculate the mean and variance of the original distributions and compare to the derived (summed) series.