Prof. Dr. Hauke Schramm
Department of Computer Science
Kiel University

Pattern Recognition

Exercises

Practice Sheet 1

Exercise L-1.1 (Octave basics, random variables and statistics)

a. Generate two N x N matrices `A` and `B` containing integer random values (set `N = 100`). The numbers in `A` must be uniformly distributed in the interval [0,10) and in `B` normally distributed with mean 0 and standard deviation $\sigma = 10$. Utilize the Octave functions `rand`, `randn`, `fix` and `round`. For the generation of `A` you need `fix` and for `B` `round`. Why?

b. Determine the minimum and maximum element of `A` and `B` (functions min and max). Store the results in `amin`, `amax`, `bmin` and `bmax`.

c. Determine the frequency of each matrix element in `A` (respectively `B`) and store it as follows in vector `a` (resp. `b`): The first element of `a` (`b`) contains the frequency of the smallest element in `A` (`B`) and the last element of `a` (`b`) contains the frequency of the largest element in `A` (`B`).

   Hints:

   i. For this task you may use the functions `find` and `length` and work with a `for`-loop.
   ii. `A == i` returns a N x N logical matrix with values 1 only at elements in `A` which equal `i`.
   iii. You may append a number `x` as follows to a vector `a`: `a = [a  x];`

d. Draw the frequency distributions in vector `a` and `b` into two separate figures (Figure 1 and Figure 2) using the functions `bar` and `figure`. Label them appropriately. Note that this graphical representation of the distribution of data is called a ***histogram***.

   Useful commands: `title`, `xlabel`, `ylabel`, `axis`

e. Since the data in matrix `A` is uniformly distributed, all components of vector `a` should have similar values which must be evenly distributed around the mean frequency value. Illustrate this fact by plotting the mean of all components in `a` as horizontal line into Figure 1. Use the command "`hold on`" to retain the current figure when plotting a new object.

   Useful commands: `line, hold on`

f. Generate probability distributions from the histograms using relative frequencies and plot them into Figure 3 (vector `a`) and 4 (vector `b`). Show that the probabilities in each distribution sum up to 1.

g. The data in matrix `B` is normally distributed with mean value $\mu = 0$ and standard deviation $\sigma = 10$. Its probability in Figure 3 can be described by the following function:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

   Draw this function as red line into Figure 4, using for the abscissa a stepsize of 0.1 between the end points bmin and bmax.

h. Check the following statements in your program:

   a. 68.3% of all elements in B lie in the interval [$-\sigma$, σ]
   b. 95.5% of all elements in B lie in the interval [-2σ, 2σ]
   c. 99.7% of all elements in B lie in the interval [-3σ, 3σ].

Prof. Dr. Hauke Schramm
Department of Computer Science
Kiel University

Exercise L-2.1 (Octave basics, random variables and statistics)

Suppose we consider two random variables:

- x: Weather condition, where x=1 represents sunny weather and x=2 represents rainy weather.
- y: Driving behavior, where y=1 represents cautious driving and y=2 represents fast driving.

The joint probabilities are defined as follows:

- P(x=1,y=1)=0.2: The probability that the weather is sunny and driving is cautious.
- P(x=1,y=2)=0.3: The probability that the weather is sunny and driving is fast.
- P(x=2,y=1)=0.4: The probability that the weather is rainy and driving is cautious.
- P(x=2,y=2)=0.1: The probability that the weather is rainy and driving is fast.

For all other combinations, the probability P(x,y)=0.

This joint distribution can be represented as a matrix:

	$Y=1$	$Y=2$
$X=1$	0.2	0.3
$X=2$	0.4	0.1

Tasks

1. Calculate the marginal probabilities P(x=1) and P(x=2).
2. Calculate the marginal probabilities P(y=1) and P(y=2).
3. Verify if the sum of all joint probabilities is equal to 1.

Solve this using an Octave script.