Computational Intelligence, SS11

http://www.spsc.tugraz.at/courses/computational-intelligence

Paul Meissner [1]

5 Homework: Maximum Likelihood Estimation and Bayesian Classification

[Points: 17+8* , Issued: 2011/06/10 , Deadline: 2011/06/24 , Tutor: Georg Kapeller [2]; Infohour: TBA , Room and Date will be posted in the newsgroup , ;

5.1 Maximum Likelihood Estimation and Classification [17 Points]

General: In this task, we want to classify so-called line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios in wireless communication systems. In a NLOS situation, the direct propagation path between transmitter and receiver is blocked, which can cause significant difficulties. Imagine an indoor radio communication scenario where a transmitter sends out a carrier signal (i.e., a sine-wave). The signal at the receiver consists of a sum of many scaled and delayed waves, due to reflections of the signal at walls or other obstacles and other effects like scattering. As this is rather complicated to describe mathematically, the amplitude and phase of this carrier at the receiver is usually characterized statistically.

It is well known that the distributions of amplitude and phase differ for LOS and NLOS cases, which we want to use to build a statistical classifier. We denote the amplitude as feature x_1 and the phase as x_2. In LOS situations, the amplitude of the received carrier follows a Ricean distribution, i.e.

$$p(x_1|t = \mathrm{LOS}) = \frac{x_1}{\sigma^2} e^{\frac{-(x_1^2+\nu^2)}{2\sigma^2}} I_0\left(\frac{x_1 \nu}{\sigma^2}\right) \tag{1}$$

where $I_0(\cdot)$ is a Bessel function. In NLOS situations, the amplitude follows a Rayleigh distribution

$$p(x_1|t = \mathrm{NLOS}) = \frac{x_1}{\sigma^2} e^{\frac{-x_1^2}{2\sigma^2}}. \tag{2}$$

The phase of the carrier can be modeled as a Gaussian random variable in LOS situations, i.e. $p(x_2|t = \mathrm{LOS}) = \mathcal{N}(x_2|\mu_{x_2}, \sigma_{x_2}^2)$ and follows a uniform distribution between $-\pi$ and π in NLOS situations, i.e. $p(x_2|t = \mathrm{NLOS}) = \mathcal{U}(x_2|[-\pi, \pi]) = \frac{1}{2\pi}$.

In the following, we want to design an ML and a Bayesian classifier that decides either based on x_1 only, or on both x_1 and x_2.

1. Download the given data set from the course homepage and load it into the Matlab workspace. The file `data_5_1.mat` is to be used in this task, it contains training and test data, both split up into the LOS and NLOS classes. Feature x_1, the amplitude, is the first row of the data sets, feature x_2, the phase, is the second row of the data sets.

2. Make a scatter plot of the training data (i.e., plot x_2 over x_1) and comment your expectations in the classification performance, either based on x_1 only or on both x_1 and x_2!

[1] mailto:paul.meissner@tugraz.at
[2] mailto:georg.kapeller@student.tugraz.at

5.1.1 Classification using the amplitude only

1. In this subtask, we classify only using the feature x_1. To this end, we need to estimate the likelihood models, i.e., the parameters of the PDFs in equations (1) and (2). First, calculate the maximum likelihood estimate of the Rayleigh parameter σ based on a list of N i.i.d. samples analytically. Implement the obtained estimator for the NLOS training data.

2. For the Rice distribution, no closed form ML estimator exists. Hence we use a kernel density estimator. For this you can use the Matlab function `[rice_pdf, x_rice] = ksdensity(...)` on the LOS training data. To evaluate the estimated Ricean PDF at a new sample, say x_n, you can use interpolation: `p_xn = interp1(x_rice, rice_pdf, x_n)`.

3. To evaluate your estimators graphically, plot both the estimated Rayleigh and Rice PDFs over the two class histograms (use `[N,X] = hist(...)` in Matlab) in one plot. Therefore, you will need to normalize the histograms. How is this normalization done?

4. Also estimate the prior probabilities of the two classes. Finally, perform both ML and Bayesian classification using the test data. Report and comment the classification performance and make scatter plots that show the classification results! How well can you do using only the amplitude?

5. In which ways are the results of ML and Bayesian classifiers different in this example? Does the better overall performance of one of the classifiers come with a cost?

5.1.2 Classification using both amplitude and phase

1. Now we also take into account the feature x_2. You can assume that x_1 and x_2 are statistically independent. Again, find the likelihood models for x_2 and plot them over the corresponding normalized histograms.

2. Perform ML and Bayesian classification using the test data! Again make scatter plots that show the classification results and comment them.

5.1.3 Performance evaluation of the ML estimator [8* Points]

In this task, we want to evaluate the performance of the maximum likelihood estimator of the Rayleigh parameter σ from data sets of various sizes.

1. Realizations of a Rayleigh distributed random variable can be generated as follows: Assume we have two independent, zero-mean Gaussian random variables X and Y with variance σ^2. Then, the random variable $Z = \sqrt{X^2 + Y^2}$ follows a Rayleigh distribution with parameter σ. In the following, we estimate σ from a list of N samples of Z using the ML estimator from the previous task.

2. Generate samples of Z with $\sigma = 2$. For different data set lengths N up to 2000, estimate σ and plot true value and the estimated values over N. What can you say about the variance of the estimator as N increases?

3. Generate samples of Z for $N \in \{100, 1000, 10000\}$ and estimate σ, but now repeat the process 10000 times for each N. Make sure that you use a new set of samples for each estimation! For each of the three data set lengths N, plot a histogram of the estimate of σ. What distribution do you observe for the ML estimator and what is the influence of N?