Exercise 1

In density estimation with Parzen windows, we distribute hypercubes of fixed size in the dataspace and the number of samples involved in these hypercubes is estimated based on the given data distribution. Assume that an 1D dataset is observed and the number of samples per hypercube is given by $K = \sum_{i=1}^{n} k(\frac{x-x^{(i)}}{h})$ where $k(u) = \begin{cases} 1 & |u| \leq 1/2 \\ 0 & otherwise \end{cases}$

a) Prove that the pdf of the data distribution is given by $p(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h} \cdot k(\frac{x-x^{(i)}}{h})$.

b) Analyze the effect that the window width has on the estimation of the function $p(x)$.

Exercise 2

Consider a histogram-like density model in which the space x is divided into fixed regions for which the density $p(x)$ takes the constant value h_k over the k^{th} region, and that the volume of region k is denoted Δ_k . Suppose we have a set of n observations of x such that n_k of these observations fall in region k. Using a Lagrange multiplier to enforce the normalization constraint on the density, derive an expression for the maximum likelihood estimator for the h_k.

Exercise 3

Show that the K-nearest-neighbour density model defines an improper distribution whose integral over all space is divergent.

Exercise 4

Suppose that for an image recognition task, we want to extract the principal components from a set of n images on a mobile robot. We reshaped each image into a column vector and now we have $d \times n$-dimensional data X which has n samples of $d \times 1$-dimensional vectors. These vectors correspond to image data and the dimension $d \gg n$. The principal components correpond to eigenvector of covariance matrix of C

$$C = \frac{1}{n-1} X_c {X_c}^\top \tag{1}$$

where X_c is obtained by subtracting the mean vector from X. Since for a large value of d (for an image of size 640×480, $d = 307200$), this computation can easily hang the On-board computer. A useful trick is to calculate the eigenvector of C_1 which is $n \times n$

$$C_1 = \frac{1}{n-1} {X_c}^\top X_c \tag{2}$$

and now if v_1 is an eigenvector of C_1 with corresponding eigenvalue λ then $v = X_c v_1$ is an eigenvector of C with corresponding eigenvalue λ. Show that this claim is true.

Exercise 5

Use prove by induction to show that the linear projection onto an M-dimensional subspace that maximizes the variance of the projected data is defined by the M eigenvectors of the data covariance matrix S, corresponding to the M largest eigenvalues.