Machine Learning WS2015/16

Last Homework Assignments

Task 1: Law of total covariance (2 points)

Prove the following identity: $Cov[\mathbf{x}] = E[Cov[\mathbf{x}|\mathbf{s}] + Cov[E[\mathbf{x}|\mathbf{s}]]$

where $Cov[\mathbf{u}] = E[\mathbf{u}\mathbf{u}^\top] - E[\mathbf{u}]E[\mathbf{u}]^\top$ and $Cov[\mathbf{u}|\mathbf{v}] = E[\mathbf{u}\mathbf{u}^\top|\mathbf{v}] - E[\mathbf{u}|\mathbf{v}]E[\mathbf{u}|\mathbf{v}]^\top$.

Task 2: MNIST & PCA (3 points)

Download the file *mnist_train.mat* from Ilias and perform the following processing steps for the training data of the two classes ($X0 = train\{10\}; X1 = train\{1\};$):

a) Compute the two class conditional means and variances and visualize each of them as an image.

b) Center the data by subtraction of the global mean (i.e. not the class conditional means) and perform a principal component analysis (PCA). Determine the number of dimensions m that are necessary to preserve 90% of the energy (i.e. find the minimal m for which $\frac{\sum_{k=1}^{m} D_{kk}}{\sum_{k=1}^{784} D_{kk}} > 0.9$) where D_{kk} denotes the k-th eigenvalue. Define U_m as the matrix containing the first m columns of the eigenvector matrix and left-multiply with $U_m^\top$ to obtain new data matrices $Y0, Y1$ that contain only m dimensions for each image. Determine a reconstruction of the original data matrices $\hat{X}0, \hat{X}1$ by left-multiplying $Y0, Y1$ with U_m and compute the relative average reconstruction error $E[||\mathbf{x} - \hat{\mathbf{x}}||^2]/E[||\mathbf{x}||^2]$.

c) Generate new data matrices $Z0, Z1$ that contain only the first two principal components of the data by left multiplying with $U_2^\top$. Determine the fraction of energy $\frac{\sum_{k=1}^{2} D_{kk}}{\sum_{k=1}^{784} D_{kk}}$ preserved by these two components.

Task 3: MNIST classification (9 points)

Continue Task 2 as follows:

a) Use the dimensionality reduced data matrices $Z0, Z1$ and make a plot that shows the samples from each class as dots in two different colors.

b) Fit a Gaussian distribution to each of the datasets $Z0$ and $Z1$. Plot the decision boundary of the MAP estimator obtained with these two Gaussian likelihoods when assuming equal prior probabilities for the two classes.

c) Download the file *mnist_test.mat* from Ilias and compute the average $0-1$-loss $E[|c - \hat{c}_{MAP}|]$ of the MAP classifier $\hat{c}_{MAP}(\mathbf{z}) \in \{0, 1\}$ on the training and on the test set where $c \in \{0, 1\}$ denotes the true class label.

d) Build a classifier

$$\hat{c}_{lin}(\mathbf{z}) = \operatorname{sgn}\left(\underbrace{E[c\mathbf{z}^\top](E[\mathbf{z}\mathbf{z}^\top])^{-1}}_{\mathbf{w}^\top}\mathbf{z} - \vartheta\right)$$

by first computing a linear regression to find an informative linear projection $\mathbf{w}^\top$ (a "linear discriminant") and then finding an optimal threshold ϑ. Plot the decision boundary $\hat{c}_{lin}(\mathbf{z}) = 0$ into the same plot as before and compute its average $0-1$-loss on the training and on test set.

e) Perform a linear and quadratic feature space embedding $\mathbf{\Phi} := (x_1, x_2, x_1^2, x_2^2, x_1x_2)^\top$ and then do the same as in (d) but using $\mathbf{\Phi}$ instead of $\mathbf{z}$ as input.