LEARNING ALGORITHMS AND SYSTEMS LABORATORY
FACULTE DES SCIENCES ET TECHNIQUES DE L'INGENIEUR
SCHOOL OF ENGINEERING / SWISS INSTITUTE OF TECHNOLOGY

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Lecturer:
Prof. ***Aude Billard***
Office: *ME A3 464*
direct: *+41-21-693 54 64*
secretary:+41-21-693 09 39
fax: *+41-21-693 78 50*
E-mail: aude.billard@epfl.ch
Web: http://lasa.epfl.ch/

Assistant:
Dr. **Basilio Noris**
Office: *ME A3 395*
direct: *+41-21-693 7824*
E-mail: basilio.noris@epfl.ch

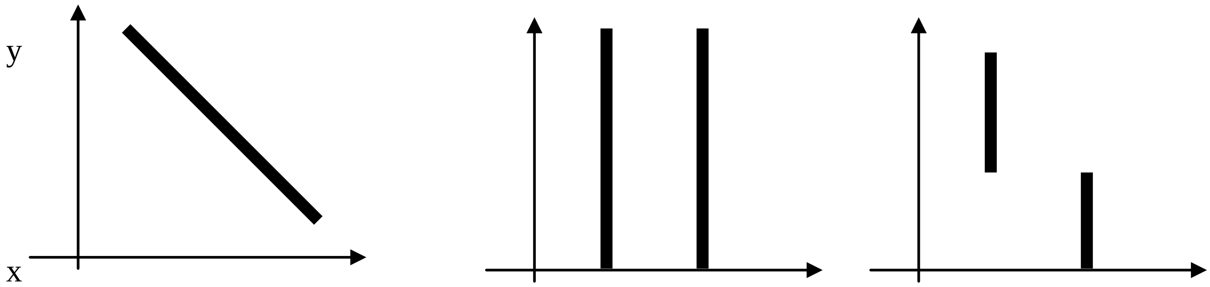
Applied Machine Learning

Exercises - I

WINTER 2011-2012

<u>Exercise 1: Principal Component Analysis</u>

a) For which of the following distributions are the two variables x and y statistically independent?



b) Imagine that you have run PCA on data gathered from one of the questionnaires gathered by the car manufacturer in which 10’000 people gave their age, gender, country of residence and the car model they purchased. How would you interpret the results if:

i) You find that all N eigenvectors (N=dimensionality of the dataset) covers the same % of the data variance.

ii) 1 single eigenvector covers 99% of the data variance (5%).

Exercise 2: Statistical Independence and uncorrelatedness

i) Consider two variables $x = [x_1 \ x_2] \in \mathbb{R}^2$ and $y \in \mathbb{R}$ with joint distribution $p_{x,y}(x, y)$.
Explain which of the following two joint distributions are statistically independent and why:

$p_{x,y}(x, y) = (1 - x_1)(x_1 - y)$

$p_{x,y}(x, y) = (x_1 + 3x_2) y$

ii) If x_1 and x_2 are two uncorrelated variables, show that if g and h are two linear functions, then $y_1 = g(x_1)$ and $y_2 = h(x_2)$ are still uncorrelated.
Show that if g and h are non linear, integrable function, y_1 and y_2 are uncorrelated if x_1 and x_2 are statistically independent.

iii) Show that when a N-dim. set of data points X is projected onto the eigenvectors $V=[v_1 ... v_N]$ of its covariance matrix $C=XX^T$, the covariance matrix YY^T of the projected data Y is diagonal and hence that, in the space of the eigenvector decomposition, the distribution of X is uncorrelated.

Exercise 3: PCA – Derivation of the estimation of the PCs

While in class we saw one way to compute the principal components of a distribution in one go, often it is practical to do it in an iterative manner. For this, one may iteratively minimize a functional. Following this idea,

a) Show that maximizing the functional $J(e) = e^T C e$, under the constraint $\|e\| = 1$, where *C* is the covariance matrix and *e* a vector, finds an eigenvector of *C*.

b) Further, show that if *x* has a zero mean distribution, then the distribution of $y = e^T x$ the projection of *x* on the eigenvector has zero mean.

Supplementary Exercises (to be done at home)

Exercise 1: Probabilities and Statistics

For what constant k is f(x) = k e^{-x} a probability density function on [0,1]?

Exercise 2: Principal Component Analysis

Compute the principal components and eigenvalues of the following matrix:

$$A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$$

Exercise 3: Covariance, correlation

Show that $-1 \leq corr\left(x, y\right) \leq 1$. Recall that $\mathrm{var}\left(\frac{x}{\sigma_x} + \frac{y}{\sigma_y}\right) \geq 0$ and that $\mathrm{var}(x + y) = \sigma_x + \sigma_y - 2\,\mathrm{cov}(x, y)$ and that $corr\left(x, y\right) = \frac{\mathrm{cov}\left(x, y\right)}{\sigma_x \sigma_y}$.