T-61.3050 Machine Learning: Basic Principles

Prerequisite Knowledge Test 2013

These problems test your prerequisite knowledge. This test can be submitted to the course mailing list. It does not affect passing the course or your grade in any way, but you may use it to check your knowledge before the course.

Please submit your answer electronically as a .PDF or .DOC file. Submit your answer by email to t613050@ics.aalto.fi, write "Prerequisite knowledge test" followed by your student number as the title of the email.

1 Algebra, Probabilities

Let Ω be a finite set of all possible outcomes and $P : \Omega \to \mathbb{R}$ a probability measure that (by definition) satisfies $P(\omega) \geq 0$ for all $\omega \in \Omega$ and $\sum_{\omega \in \Omega} P(\omega) = 1$. Let $f : \Omega \to \mathbb{R}$ be an arbitrary function on Ω. Define the expectation of f by $E[f(\omega)] = \sum_{\omega \in \Omega} P(\omega) f(\omega)$. The *variance* of f is defined by $\mathrm{Var}[f(\omega)] = E\left[(f(\omega) - E[f(\omega)])^2\right]$. Using these definitions, show that:

(a) $E[\cdot]$ is a linear operator. See the definition of a linear operator at http://mathworld.wolfram.com/LinearOperator.html .

(b) The variance can also be written as $\mathrm{Var}[f(\omega)] = E\left[f(\omega)^2\right] - E[f(\omega)]^2$. Hint: the proof is short if you use linearity.

2 Matrix Calculus

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix with distinct eigenvalues $\lambda_i \in \mathbb{R}$, $i \in 1, \ldots, n$, defined by $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$, where $\mathbf{v}_i \in \mathbb{R}^n$ and the eigenvectors satisfy $\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}$; where the Kronecker delta is $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ otherwise. Let $\mathbf{B} = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T$ (spectral decomposition). Show that matrix $\mathbf{B}$ has the same eigenvectors and eigenvalues as matrix $\mathbf{A}$ (if you want, you can also show that $\mathbf{A} = \mathbf{B}$; this is however not required here).

3 Algorithms

The Fibonacci numbers $F(i)$ are defined for $i \in \mathbb{N}$ recursively as $F(i + 2) = F(i + 1) + F(i)$, with $F(1) = F(2) = 1$. Using pseudo-code, write down an algorithm that takes $n \in \mathbb{N}$ as an input and outputs the Fibonacci numbers from $F(1)$ to $F(n)$. Analyze the time complexity of your algorithm using the O-notation. What can you say about efficiency of your algorithm?

4 Basic Data Analysis, Software Tools

A mystery data set at https://noppa.tkk.fi/noppa/kurssi/t-61.3050/harjoitustyot/T-61_3050_data_set.txt has 4000 data items (rows), each having 16 real-valued variables (columns). Write a small program in R, S, Octave or Matlab that loads the data set, finds the two variables having the largest variances and makes a scatterplot of the data items using these two variables. Attach a printout of your program code and the scatterplot produced by your program as an answer to this problem. Note that the program should choose the variables without user interaction.