

Preprocessing in Data Mining. The Rudimentary Operations in R. Laboratory 1

Krystian Jobczyk

October 1, 2017

0.1 (Manipulation on data represented by vectors).

Consider the sequence of (rational) numbers of the length = 10 from -1 to -2.
Write the appropriate R-code that:

- A generates this sequence,
- B writes in the third element of the sequence,
- C writes in all element without the first and the fifth one.
- D writes in each value divided by the mean of all elements,
- E ensures that the 8th element of the sequence from A is equal to 7th element
of the sequence defined in point D.
- F ensures that two first elements of a vector from A are equal to zero and
values from the positions 4 and 5 are exchanged by values 7,8.

0.2 (Diagrams of functions).

Writhe the appropriate R-code to obtain the diagrams of the following functions:

A $f(x) = \frac{(\arctan(x) + \log_2(x))}{xe^x + 5\cos(x)}$, for $x \in [-10, 0]$,

B $f(x) = \lim_{n \rightarrow \infty} \frac{1 + \sin(x) + \sin^2(x) + \dots + \sin^n(x) + \dots}{1 + \tan(x) + \tan^2(x) + \dots + \tan^n(x) + \dots}$, for $x \in [-\pi, \pi]$,

C $f(x) = \lim_{n \rightarrow \infty} \arcsin(\sqrt[n]{2^{nx} + 3^{nx} + 4^{nx}})$, for $x \in [1, 2]$,

D $f(x) = \arctan\left(\frac{x+2x+3x+\dots+kx}{\sqrt[k]{1+x} + \sqrt[k]{2+k} + \dots + \sqrt[k]{k+x}}\right)$, for $x \in [0, 1)$), $k \rightarrow \infty$.

E $f(x) = \frac{6}{\pi^2} \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \right) x^2 - \left(\sqrt{2^{\ln e^2}} \right) x + 1$, for $x \in [-1, 1]$.

$F^* f(x) = \binom{m}{1} \cos^{m-1} x \sin x - \binom{m}{3} \cos^{m-3} \sin^3 x + \binom{m}{5} \cos^{m-5} \sin^5 x - \binom{m}{7} \cos^{m-7} \sin^7 x + \dots + \binom{m}{0} \cos^m x - \binom{m}{2} \cos^{m-2} \sin^2 x + \binom{m}{4} \cos^{m-4} \sin^4 x - \binom{m}{6} \cos^{m-6} \sin^6 x,$
 for $m = 2017$, $x \in [-2\pi, 2\pi]$. (**Hint:** Use De Moivre's identity for complex numbers).

0.3 (Manipulation on data represented by matrices).

Define two matrices: the matrix A – containing natural numbers from 1 to 9 – and such a matrix B that its first column contains as the values:

- the number of 3-elemental combination of the 4-elemental set,
- logarithm from the 4-elemental combination of the 5-elemental set,
- the logarithm from the factorial of 3.

The values of the second column are the initial Fibonacci numbers and the values of the third one – are the initial prime numbers. (respectively). Basing on the appropriate R-code:

- A Find the matrices $A \bullet B$ and $B \bullet A$, where \bullet denotes the operation of matrix multiplication (why \bullet is feasible?),
- B Find the eigenvalues and eigenvectors of $A \bullet B$ and $B \bullet A$,
- C Find the traces of both matrices.

0.4 (R as an interactive programming language).

Define the temperature converters in R in terms of the Fahrenheit, Celsius and Kelvin scales, if the mutual relations among these temperature scales may be specified as follows:

- $T_{Celsius} = T_{Kelvin} - 273, 15$ and
- $T_{Fahrenheit} = \frac{9}{5}T_{Celsius} + 32$.

Consider the of temperature spectrum from -30 to 25 degrees of Celsius.