CSC433 Scripting for Data Analysis Assignment 6

Matt Triano

May 27, 2017

This assignment is intended to show the performance difference between using R's vectorized methods and iterative approaches common to many other programming styles. As the graphs below show, R's vectorized methods are far faster than the iterative approach.

From the plot of runtimes for both functions, we see that runtimes for the mycbind() function increases according to a polynomial trend, and this trend dominates the runtimes for the regular cbind() function. From the plot scaled to just the runtimes for the cbind() function, we see that runtimes do increase as n increases, but at least orders of magnitude more slowly than the runtimes for the mycbind function.

In conclusion, when developing R scripts for large datasets, vectorized methods should be preferred over iterative coding approaches.

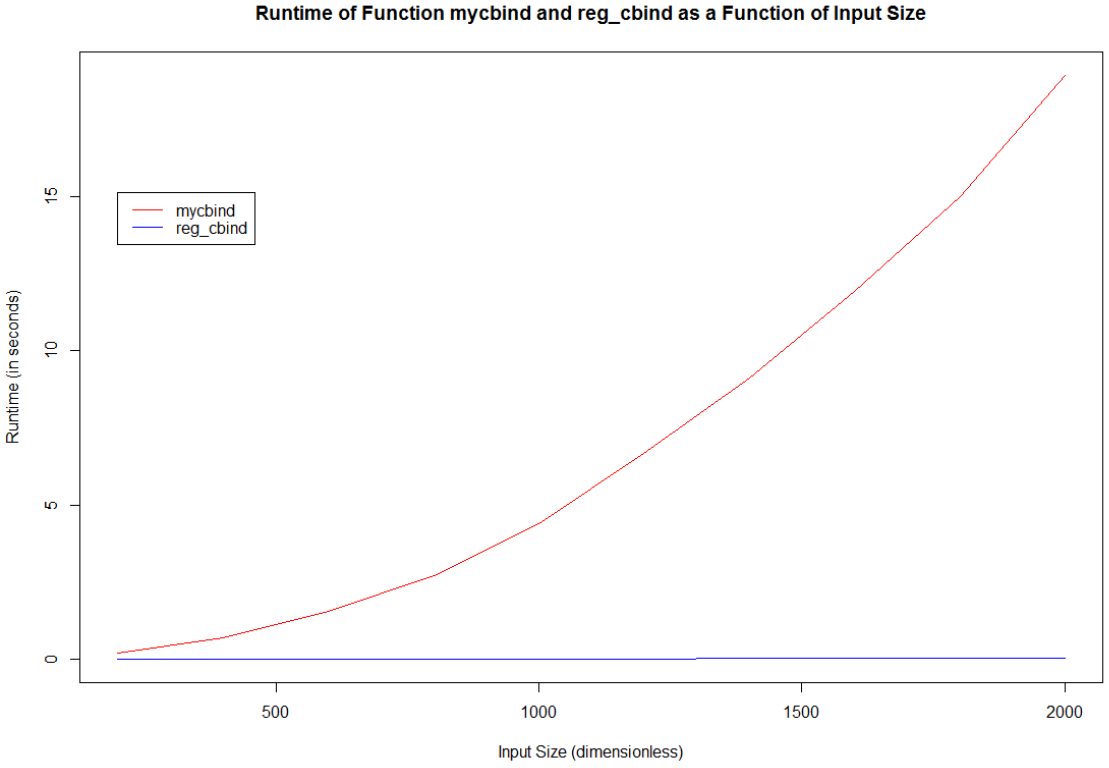

Figure 1: Runtimes for Both mycbind() and regular cbind()

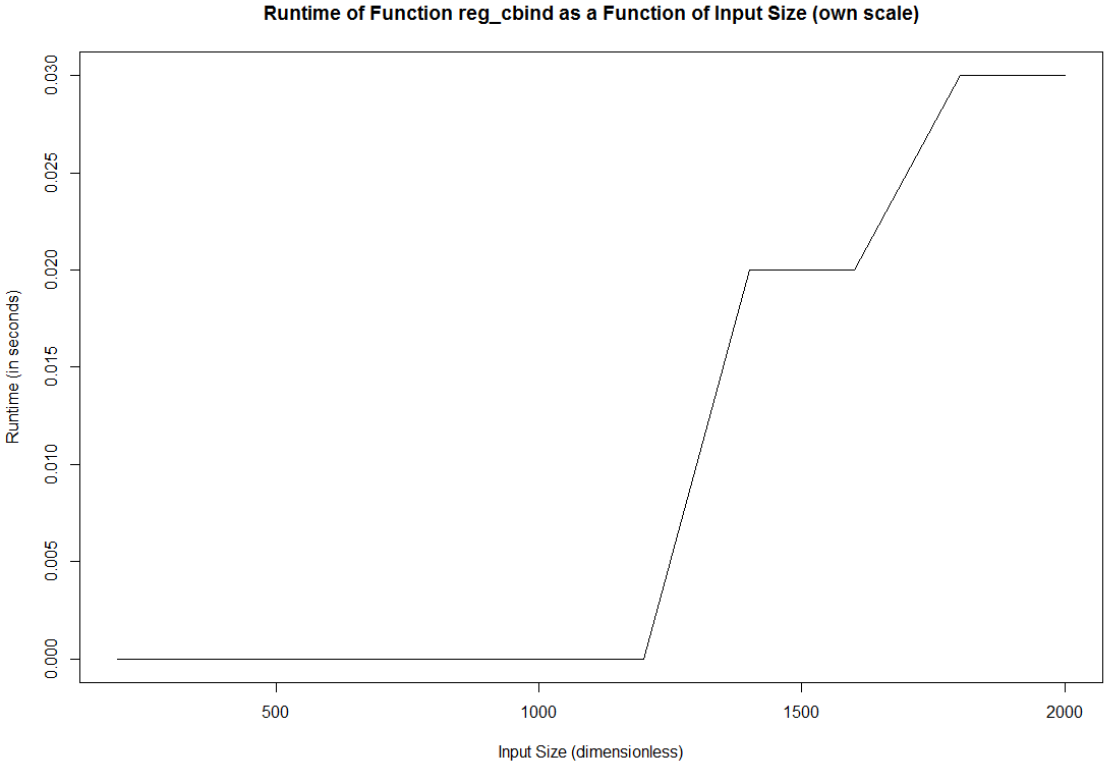

Figure 2: Runtimes for regular cbind()