A first glimpse into ‘R.ff’, a package that virtually removes R’s memory limit

Jens Oehlschlägel   Daniel Adler   Oleg Nenadic   Walter Zucchini

The availability of large atomic objects through package ‘ff’ can be used to create packages implementing statistical methods specifically addressing large data sets (like sub-bagging or package biglm). However, wouldn’t it be great if we could apply all of R’s functionality to large atomic data? Package ‘R.ff’ is an experiment to provide as much as possible of R’s basic functionality as ‘ff-methods’. We report first experiences with porting standard R functions to versions operating on ff objects and we discuss implications for package authors (and maybe also R core). Instead of a summary, here we just quicken your appetite through the list of functions and operators where we have first experimental ports:

! != \%=\%= \%*\%= \%/\%= | * + - / <= >= ^ abs acos acosh asin asinh atan atanh besselI besselJ besselK besselY beta ceiling choose colMeans colSums cos cosh crossprod cummax cummin cumprod cumsum dbeta dbinom dcauchy dchisq dexp df dgamma dgeom dhyper digamma dlnorm dlogis dnbinom dnorm dpois dsignrank dt dunif dweibull dwilcox exp expm1 factorial fivenum floor gamma gammaCody IQR is.na is.nan jitter lbeta lchoose lfactorial lgamma log log10 log1p log2 logb mad order pbeta pbinom pcauchy pchisq pexp pf pgamma pgeom phyper plnorm plogis pnbinom pnorm pois psigamma psignrank pt punif pweibull pwilcox qbeta qbinom qcauchy qchisq qexp qf qgamma qgeom qhyper qlnorm qlnorm qnorm qpois qsignrank qt quantile qunif qweibull qwilcox range range rbeta rbinom rcauchy rcishq rexp rf rgamma rgeom rhyper rlnorm rlogis rnbinom rnorm round rowMeans rowSums rpois rsignrank rt runif rweibull rwilcox sample sd sign signif sin sinh sort sqrt summary t tabulate tan tanh trigamma trunc var.