RcppCNPy: Read-Write Support for NumPy Files in R

Dirk Eddelbuettel

Debian and R Projects

Wush Wu

Institute of Electrical Engineering, National Taiwan University

26 August 2016

Summary

Python is a widely-used and popular programming language. It is deployed in use cases ranging from simple scripting to larger-scale application development. Python is also popular for quantitative and scientific application due to the existence of extension modules such as NumPy (which is shorthand for Numeric Python) and many other packages for data analysis.

NumPy (van der Walt, Colbert, and Varoquaux 2011) is used to efficiently represent N-dimensional arrays, and provides an efficient binary storage model for these files. In practice, N is often equal to two, and matrices processed or generated in Python can be stored in this form. As NumPy is popular, many project utilize this file format.

R has no dedicated reading or writing functionality for these type of files. However, Carl Rogers has provided a small C++ library called cnpy (Rogers 2015). Using the *Rcpp modules* feature in Rcpp (Eddelbuettel 2013,Eddelbuettel et al. (2016)), we provide (some) features of this library to R via the RcppCNPy package (Eddelbuettel and Wu 2016).

References

Eddelbuettel, Dirk. 2013. *Seamless R and C++ Integration with Rcpp.* Use R! New York: Springer. doi:10.1007/978-1-4614-6868-4.

Eddelbuettel, Dirk, and Wush Wu. 2016. *RcppCNPy: Read-Write Support for 'NumPy' Files via 'Rcpp'.* https://CRAN.R-Project.org/package=RcppCNPy.

Eddelbuettel, Dirk, Romain François, JJ Allaire, Kevin Ushey, Qiang Kou, John Chambers, and Douglas Bates. 2016. *Rcpp: Seamless R and C++ Integration.* https://CRAN.R-Project.org/package=Rcpp.

Rogers, Carl. 2015. "Cnpy." *GitHub Repository.* https://github.com/rogersce/cnpy; GitHub.

van der Walt, Stéfan, S. Chris Colbert, and Gaël Varoquaux. 2011. "The NumPy Array: A Structure for Efficient Numerical Computation." *Computing in Science & Engineering* 13 (2): 22–30.