Sparse Matrices in package Matrix and applications

Martin Mächler[1,2,*], **Douglas Bates**[1,2]

1. ETH Zurich, Switzerland and University of Wisconsin, Madison, USA
2. R Core Development Team * Contact author: maechler@R-project.org

Keywords: sparse matrices, S4 classes and methods, large data

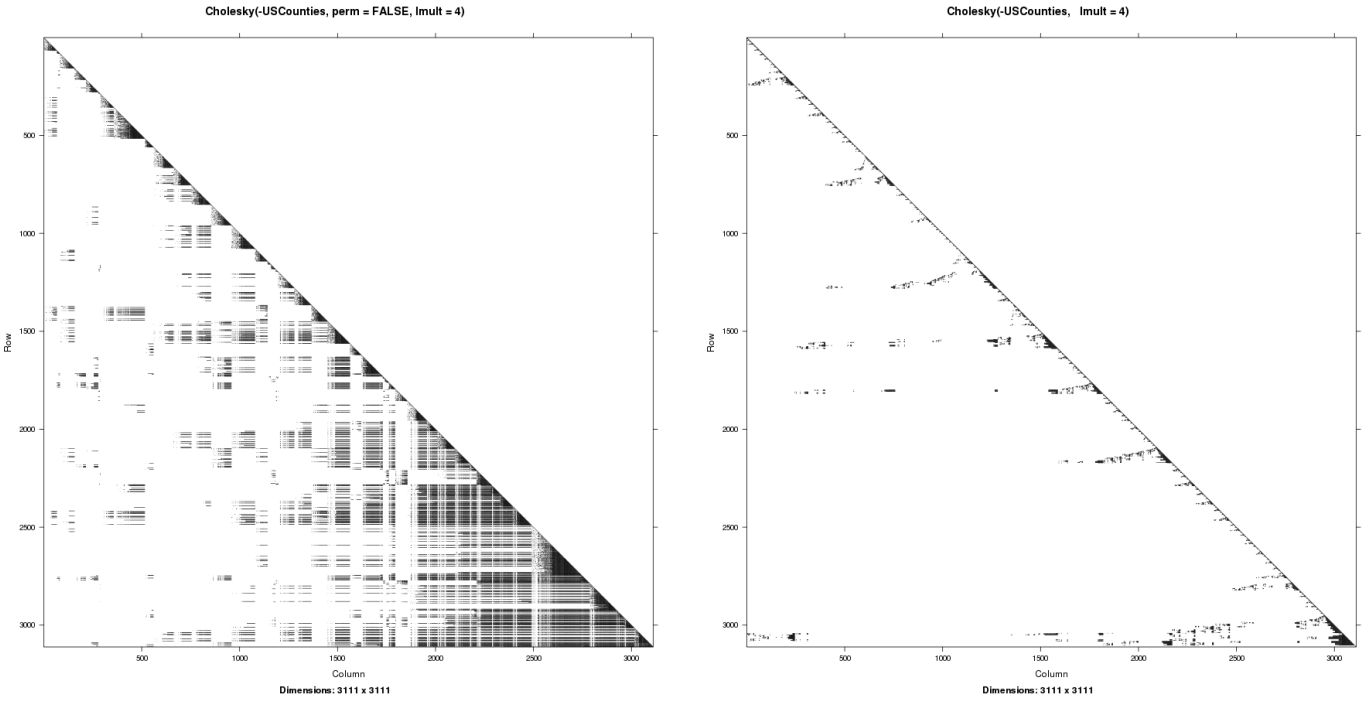
Linear algebra is at the core of many areas of statistical computing and from its inception the S language has supported numerical linear algebra via the `matrix` type and several functions and operators. However, these data types and functions do not provide direct access to all of the facilities for efficient manipulation of dense matrices, as provided by the Lapack subroutines, and they do not provide for manipulation of sparse matrices. The Matrix package provides a set of S4 classes for dense and sparse matrices that extend the basic matrix data type. Methods for a wide variety of functions and operators applied to objects from these classes provide efficient access to BLAS (Basic Linear Algebra Subroutines), Lapack (dense matrix), CHOLMOD including AMD and COLAMD and Csparse (sparse matrix) routines. The talk will focus on the *sparse* matrix classes and methods and mention some specific applications, notably in other CRAN packages.

An introduction into the sparse matrix representations, their corresponding Matrix classes, and typical constructor, inspection and visualization functions.

We will explore the space of Matrix classes, its many hundreds of methods and explain why the useR typically does not have to know most of these details.

Least squares fitting with large sparse design matrices can be accomplished via efficient sparse Cholesky decompositions and allows solving systems of sizes that would not be possible using traditional "dense" matrices.

One novel application in spatial statistics, needs to compute the determinant of $\det|I - \rho W|$ for many values of ρ and a large sparse $n \times n$ matrix W (e.g. $n = 15'000$). This problem can be reduced to compute the symbolic part of the cholesky decomposition of A *once* only, and updating the decomposition to the one of $W - \lambda I$ for varying λ, which is comparatively fast.



References

Douglas Bates and Martin Maechler (2005 ff). Introduction to the Matrix package (and other vignettes). `http://cran.r-project.org/web/packages/Matrix/`