
Subject: Re: Any interest in READ/WRITE functions for sparse (CSR) matrices?

Posted by [tom.grydeland](#) on Mon, 26 Nov 2012 13:49:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Monday, November 26, 2012 10:14:23 AM UTC+1, tom.gr...@gmail.com wrote:

> I tried searching for a bit, but couldn't find an established (binary) format for sparse matrices. Hence, I've done what one should never do -- dreamt up my own file format for saving such data.

I should perhaps point out that IDL's SPRSIN and friends would not work for us for a number of reasons, since we use the sparse format as a data container on a subset of a regular grid, and with no particular eye towards doing linear algebra on the resulting matrices (so I should probably not call them matrices at all, but only arrays). Therefore, being restricted to square matrices makes little sense. Also, the mask often applies to a number of facets or properties, which is often useful to save together in one file. It is also useful to us to have a pure index file, which can be used for saving other data (e.g. after subsequent processing) in a congruent sparse format.

A quick comparison wrt file size: a 500x500 single-precision float array randomly filled at about 10% resulted in a CSR file of ~150k and an IDL AS file of about 300k. The values themselves take up ~100k, so the overhead is about 50% for CSR, about 200% for AS in this simple case. The ratio of 1/4 for the overhead is as expected. IDL AS files save a single ULONG64 index (8 bytes) per element, while CSR saves one UINT x index (2 bytes) per element, and also one UINT x count per line. (If there are more than $2^{16}-1$ elements per line, the indices will automatically be ULONG -- 4 bytes -- instead).
