

Sparse Bidiagonalization by Householder Transformations

Gary Howell

Conventional wisdom is that Householder bidiagonalization of a sparse matrix is impractical, entailing too much fill. Numerically unstable methods are therefore used. We assert that sparse Householder bidiagonalization can be competitive in terms of storage and computational speed.

Suppose we wish to approximate a sparse matrix A by a rank k approximation A_k

$$A \approx A_k = U_k \Sigma_k V_k^T$$

where the total number of nonzero entries of U_k and V_k is at most as large as the number of nonzero entries of A . In this case, Householder reduction of A to get a the first k columns of a bidiagonal matrix is not only feasible, but has operation counts and storage requirements comparable a Lanczos method. In either case, the main computational cost is in sparse-matrix dense-vector multiplications. The required algorithm already exists in the block reduction scheme used in LAPACK. A modification combining matrix vector multiplications allows greater efficiency.

One application gaining great current attention is in data mining. A term-document matrix is represented as a sparse matrix A . Entry a_{ij} of A represents the frequency of occurrence of term i in document j . In the LSI approach to document retrieval A is approximated by a rank k matrix A_k requiring less storage.