performance

Qian

July 29, 2013

Contents

1   Dense matrix computation time on DCA

Compute 10 sigular values and vectors, 50 Lanczos iterations
    R: same algorithm using irlba package
    MADlib: use svd(. . .) for dense matrix

Matrix size	R (s)	MADlib (s)
10000 x 100	0.661	6.09
10000 x 200	0.985	6.346
10000 x 400	1.6	6.545
10000 x 800	3	7.23
10000 x 1600	5.938	8.894
10000 x 3200	11.362	11.174
10000 x 6400	26.933	15.058
20000 x 6400	51.912	23.114
40000 x 6400	104.435	39.054
80000 x 6400	167.656	76.484

2 Sparse matrix computation time on DCA

- Matrix size: 64067 x 64045

- None-zero elements: 94054 (0.00229%)

- Compute first 50 sigular values (90% of all)

- R (irlba): 20.1 s

- MADlib: `svd_block` ~ 18 min, `svd_sparse_native` ~ 445.6 sec

- MADlib reconstruction error: 2.15, relative 4.3%

3 In the above tests, R produces the exactly same results as MADlib