Randomized SVD

Brady Metherall

25 November 2019

Algorithm 1 Randomized SVD.

Input: Matrix A of size $m \times n$, Int r, Int l.
Output: Matrix of size $m \times n$ of rank-r.
$\Omega \leftarrow \mathcal{N}(0,1)^{n\times(r+l)}$
$Q,\ \_ \leftarrow qr(A\Omega)$
$U,\ \Sigma,\ V \leftarrow svd(Q^T A)$
return $(Q * U)[:, 1{:}r] * \Sigma[1{:}r, 1{:}r] * V[1{:}r, :]$

Theorem 1. *If AB^T and $B^T A$ are both symmetric matrices, then and only then can two orthogonal matrices U and V be found such that $\Sigma_A = U^T AV$, and $\Sigma_B = U^T BV$ are both diagonal matrices.*

Proof. This follows from the spectral theorem ***cite. □

Theorem 2. *The best (in the Frobenius norm) rank-r approximation to a matrix, A, is obtained by the truncated SVD, A_r.*

Proof. The following proof has been adapted from [1]. The best approximation can be found by

$$M = \underset{X\in\bowtie_r}{\arg\min}\|A - X\|_F^2, \tag{1}$$

where $\bowtie_r$ is the set of all rank-r $m \times n$ matrices. The minimum error is then

$$\begin{aligned}\|A - M\|_F^2 &= \langle A, A\rangle - 2\langle A, M\rangle + \langle M, M\rangle, &(2)\\ &= \langle A, A\rangle - 2\big\langle A, U\Sigma_M V^T\big\rangle &(3)\\ &\qquad + \langle \Sigma_M, \Sigma_M\rangle.\end{aligned}$$

At the minimum, the change in $\|A - X\|_F^2$ is zero for some change in X. This change in X can be encapsulated as $U \mapsto sU$ where s is infinitesimal and antisymmetric to maintain orthogonality. Thus, at the minimum

$$0 = \langle A, sM\rangle = \big\langle AM^T, s\big\rangle. \tag{4}$$

Therefore, it is the case that AM^T is symmetric. By following a similar procedure, we find that $M^T A$ must be symmetric as well.

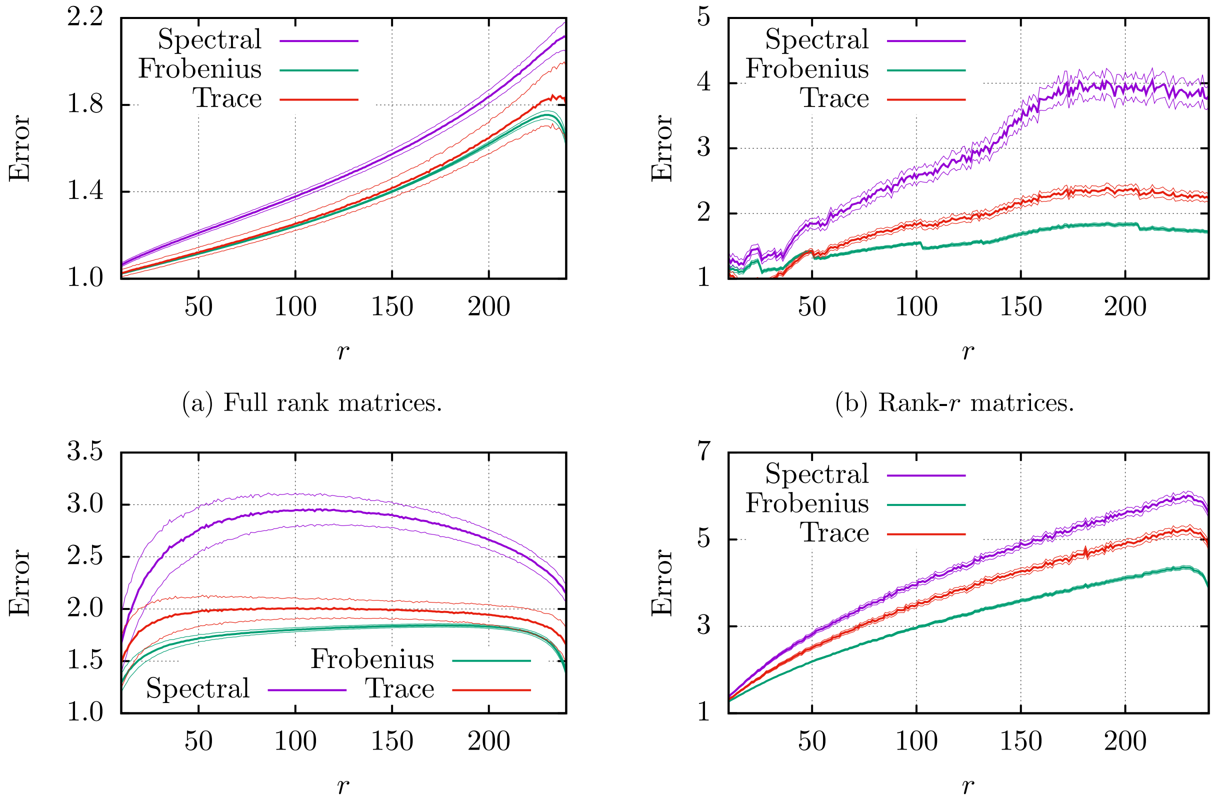
By Theorem 1 A and M exhibit the same U and V in their SVDs. Now (2) can be simplified to

$$\begin{aligned}\|A - M\|_F^2 &= \|\Sigma_A - \Sigma_M\|_F^2, &(5)\\ &= \sum_{i=1}^{n}(\sigma_i(A) - \sigma_i(M))^2, &(6)\\ &= \sum_{i=r+1}^{n}\sigma_i(A)^2. &(7)\end{aligned}$$

This minimum is indeed achieved by the truncated SVD, since $\sigma_i(A_r) = \sigma_i(A)H(r - i)$. □

References

[1] C. Eckart and G. Young, “The Approximation of One Matrix by Another of Lower Rank,” *Psychometrika*, vol. 1, pp. 211–218, Sept 1936.



(a) Full rank matrices.

(b) Rank-r matrices.

(c) Algebraically decaying singular values.

(d) Geometrically decaying singular values.

Figure 1: Relative error of four types of matrix for the spectral, Frobenius, and trace norms. The thin lines represent the standard deviation in Figures 1a and 1c, and the standard error in Figures 1b and 1d since the calculations are less numerically stable.