International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.24, July 2025

The Critical Role of Hyperparameter Tuning in Machine Learning: A Focus on the SVD Method for Matrix Completion

Faezeh Aghamohammadi
Department of Applied Mathematics
Faculty of Mathematics and Computer Science
Amirkabir University of Technology (Tehran Polytechnic)
No. 350, Hafez street, Tehran, 1591634311, Iran

Fatemeh Shakeri
Department of Applied Mathematics
Faculty of Mathematics and Computer Science
Amirkabir University of Technology (Tehran Polytechnic)
No. 350, Hafez street, Tehran, 1591634311, Iran

ABSTRACT

In machine learning, the determination of hyperparameters plays an essential role. The significant impact of these parameters on the accuracy of algorithms across problem-solving scenarios cannot be denied. Improper selection of values can significantly increase errors and affects outcomes. Low rank matrix completion, an optimization problem to recover and complete a partial matrix, is an example of dealing with hyperparameter tuning. Based on the experimental knowledge, we find that establishing values for hyperparameters is imperative to achieve an optimal solution to this problem. This study investigates the hyperparameter determination of the singular value thresholding (SVT) method and proposes an approach for selecting these parameters to attain superior solutions.

Keywords

matrix completion, sampling, unique solution.

1. INTRODUCTION

In the field of machine learning, the hyperparameters tunning plays a vital and necessary role, significantly affecting the accuracy of algorithms. Inappropriate selection of these parameters often leads to a considerable increase in errors, profoundly affecting the quality of the obtained results.

Matrix completion is a problem that involves optimizing a minimization problem. Similar to many such problems, it requires tuning hyperparameters to reach the optimal solution. Most methods and algorithms proposed for matrix completion problems, such as Nuclear Norm Minimization (NNM) [3], Singular Value Thresholding (SVT) [1], Iterative Reweighted Least Square (IRLS) [4], Fixed-point iterative algorithm [8], TSNMR-based algorithm [5], Riemannian Gradient Method(RGM) [6] and OTA-MC [7] have one or more hyperparameters and each of which requires a solution to accurately determine the hyperparameter values in order to reach the optimal solution. In this article, we examine a matrix completion method called Singular Value Thresholding (SVT) [1] and propose an approach for determining its hyperparameters.

2. BACKGROUND

In general, with the technology advancement and software development, the focus on data collection and analysis has significantly increased as a crucial aspect of machine learning. An important approach for data analysis is the use of matrices or tensors. However, a significant challenge during data collection is the loss or incompleteness of the datasets. The matrix completion problem attempts to fill in missing or incomplete datasets, and several methods have been proposed to solve this problem. Assuming a low-rank structure of the target matrix, this problem is solvable and well posed [2]. By the low rank assumption, this problem is expressed as follows:

$$\min_X \operatorname{Rank}(X) \quad \text{s.t.} \quad \mathcal{P}_\Omega(X) = \mathcal{P}_\Omega(W) \tag{1}$$

Here, $W \in \mathbb{R}^{m\times n}$ is the incomplete matrix, $X \in \mathbb{R}^{m\times n}$ represents its completion, $\operatorname{Rank}(.)$ is the rank function, Ω is the set of known entries of W, and $\mathcal{P}_\Omega(.)$ is defined as [1]

$$\mathcal{P}_\Omega(A) = \begin{cases} A_{ij}, & (i,j)\in\Omega \\ 0, & (i,j)\notin\Omega \end{cases} \tag{2}$$

where A_{ij} is the $(i,\ j)$th entry of the matrix A. Since (1) is non-convex and NP-Hard [2], its convex relaxation problem is as follows [2]

$$\min_X \|X\|_* \quad \text{s.t.} \quad \mathcal{P}_\Omega(X) = \mathcal{P}_\Omega(W), \tag{3}$$

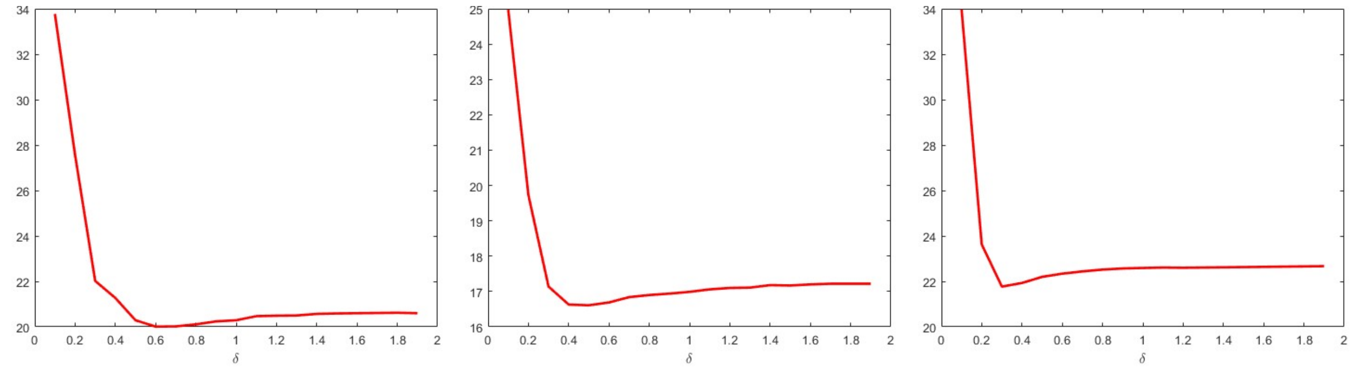
where $\|.\|_*$ signifies the sum of the singular values of the matrix.

3. MAIN IDEA

In this section, we describe the determination of the SVT method parameters, a method for solving the low rank matrix completion problem proposed in [1]. The objective function of this method is given by

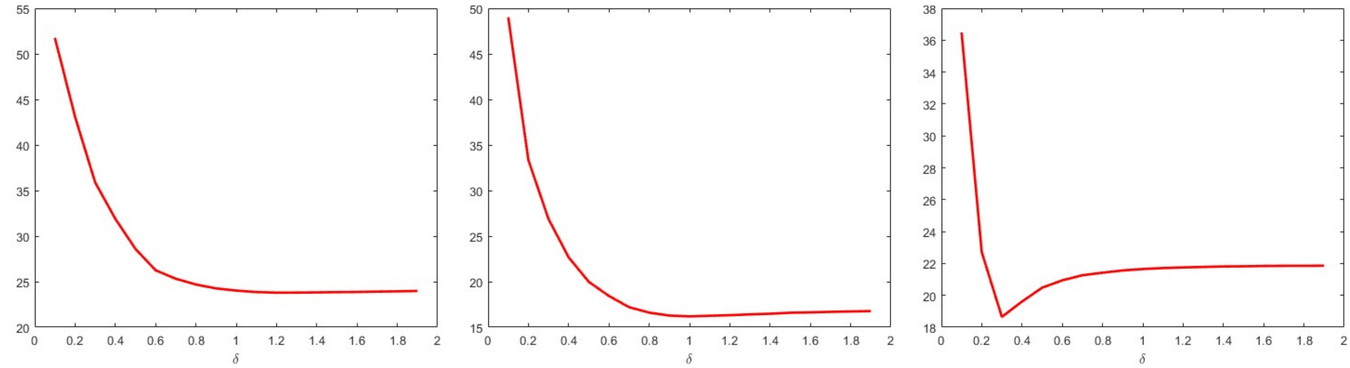
$$\min_{X,Y} \tau \|X\|_* + \frac{\delta}{2}\|X - \mathcal{P}_\Omega(Y)\|_F^2 . \tag{4}$$

X^* is the optimal solution attained through the optimization process. τ^*, δ^* and e^* denote the optimal values of τ, δ and e, which



(a) $Sr = 0.3$ (b) $Sr = 0.6$ (c) $Sr = 0.9$

Fig. 1: Plots depicting the relationship between δ and e for a 128×128 image size across different sample rates (Sr). The horizontal axis represents δ, while the vertical axis denotes e.



(a) $Sr = 0.3$ (b) $Sr = 0.6$ (c) $Sr = 0.9$

Fig. 2: Plots depicting the relationship between δ and e for a 256×256 image size across different sample rates (Sr). The horizontal axis represents δ, while the vertical axis denotes e.

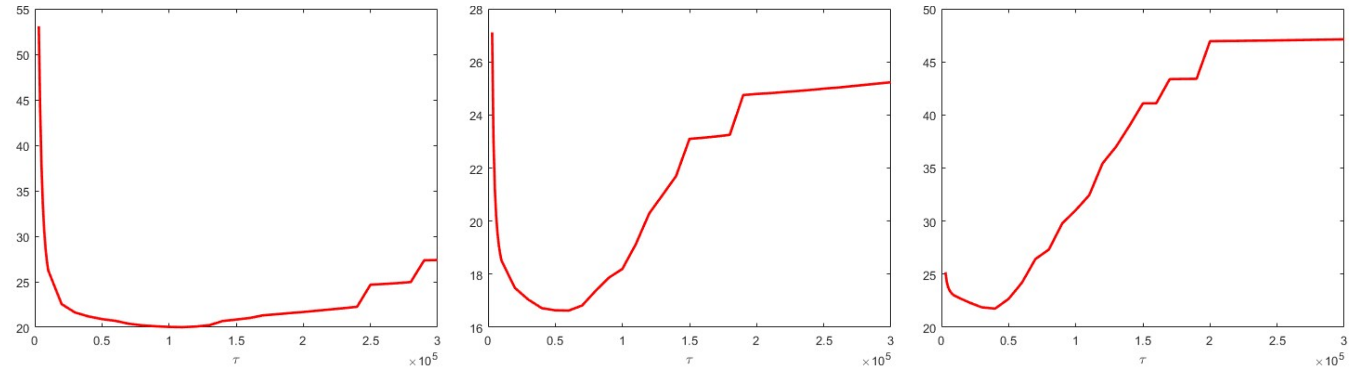
are crucial for achieving the optimal solution. The relative error, e, characterizes the algorithm's performance and is defined as follows:

$$e = \frac{\|M - X\|_F}{\|M - W\|_F} \times 100,$$

where M is the original complete matrix and the sample rate Sr is denoted as

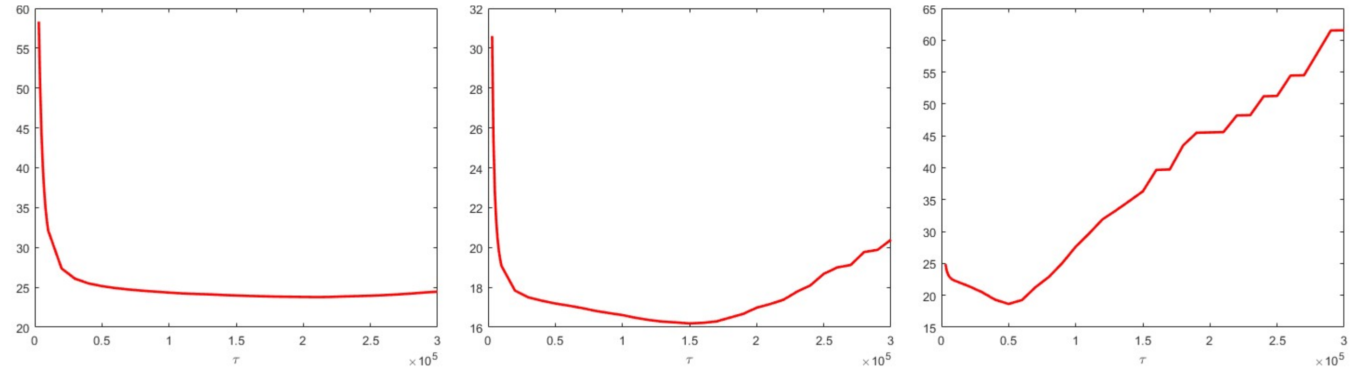
$$Sr = \frac{|\Omega|}{mn},$$

in which $|\Omega|$ is the cardinal number of sample set Ω. As is clear from (4) we must determine τ and δ to solve the problem (4).
To investigate and assess these parameters, five datasets comprising 10 black and white photos are selected, each with various dimensions: 128×128, 256×256, 512×512, 200×150, and 180×280. Using the SVT method, we experiment with different Srs ($Sr \in [0.1, 0.9]$ with a step size of 0.1). We vary the parameters τ and δ ($\tau \in [10000, 300000]$ with a step size of 10000 and $\delta \in [0.1, 1.9]$, with a step size of 0.1) for each image and observe the changes in e to determine the optimal values τ^* and δ^*.
We choose two images from the two datasets as samples to show the fluctuation in e concerning variations in the parameters δ and τ. To assess the impact of τ on the error, we set $\delta = \delta^*$. Likewise, to investigate the effect of δ on e, we set $\tau = \tau^*$.
The error changes concerning variations in δ and τ are depicted in Figures 1-2 and 3-4, respectively. Figures 1-2 illustrate a considerable increase in e when δ is inappropriately chosen. Similarly, Figures 3-4 reveal a similar issue with the τ value. Consequently, achieving the optimal solution with minimal e depends on the appropriate selection of parameters within the objective function. Therefore, these Figures emphasize the crucial role and magnitude of these parameters in problem-solving scenarios. Furthermore, Figure 5 visually displays the e changes concerning the simultaneous alterations of both τ and δ. The areas shaded in dark blue denote the lowest e values. Our objective here is to investigate and analyze the settings of two parameters τ and δ. To accomplish this goal, we use the singular value decomposition of $\mathcal{P}_\Omega(W) = U\Sigma V^T$, where, as previously mentioned, W represents the incomplete matrix and $\Sigma = \text{diag}\{\sigma_i\}_{i=1}^r$, with σ_i denoting the ith singular value of the $\mathcal{P}_\Omega(W)$ matrix, where $\sigma_1 > \sigma_2 > \ldots > \sigma_r > 0$. To determine the values related



(a) $Sr = 0.3$ (b) $Sr = 0.6$ (c) $Sr = 0.9$

Fig. 3: Plots illustrating the relationship between τ and e for a 128×128 image size across various sample rates (Sr). The horizontal axis denotes τ, while the vertical axis represents e.



(a) $Sr = 0.3$ (b) $Sr = 0.6$ (c) $Sr = 0.9$

Fig. 4: Plots illustrating the relationship between τ and e for a 256×256 image size across various sample rates (Sr). The horizontal axis denotes τ, while the vertical axis represents e.

to the δ and τ parameters, we display the correlation matrix in Figure 8. This matrix shows the Interdependence of τ and δ, as well as the dependence of τ value with Sr and the singular values of the incomplete matrix. Then, the factors influencing τ parameter encompass the sample rate (Sr) and singular values ($\sigma_i, i = 1, \ldots, r$). Based on the results of the correlation matrix, this study aims to find a formula for the parameter τ based on the two mentioned values using the experimental method. Then, we can find the δ value based on τ value using the grid search method.

Let us have a partial matrix W with $Sr = i \times 10^{-1}, i = 1, \ldots, 9$. To obtain a relation for estimating τ parameter, we initially set $\tau = \sigma_i$. However, this approach did not yield a solution close enough to the optimal solution. Subsequently, by experimenting with different coefficients and rational powers of σ_i, we discovered that the most effective solution was obtained using the following relationship:

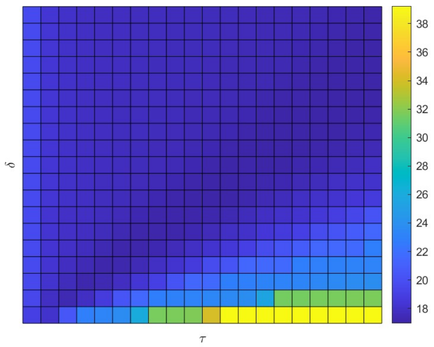
$$\tau = \sigma_i^{1.55}, \text{ s.t. } Sr = i \times 10^{-1}. \tag{5}$$

We display the error obtained using (5) with e_τ. Figures 6-7 illustrate the error variation concerning the change in $power$ in σ_i^{power} for an image from each dataset when $Sr = i \times 10^{-1}$.

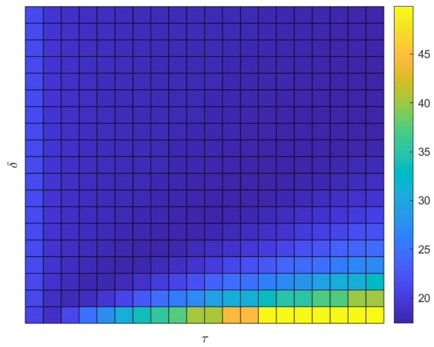
The plotted data demonstrate that the minimum error consistently appears at $power = 1.55$ in (5) across all obtained results. In addition, Figure 6 illustrates average difference between e_τ and e^* by exploring various Sr values for each of the five datasets. Table 1 shows the average $e_\tau - e^*$ for all 50 data with different Sr values. Remarkably, except for $Sr = 0.1$, the observed deviation is less than 0.2 percent, which confirms the competence of the method in obtaining the optimal solution.

4. CONCLUSION

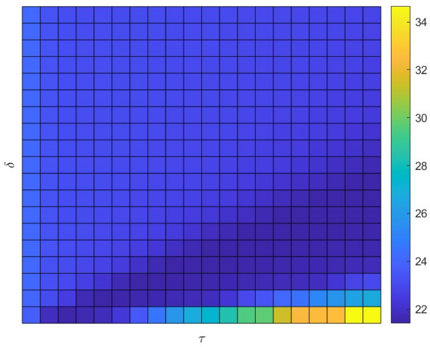
In this study, we investigated hyperparameter tuning in the singular value thresholding method to solve the matrix completion problem. The findings of this study illustrate that the proper choice of τ and δ parameters in the SVT method, as determined through the graphs and experiments, has a substantial effect on the final accuracy. Establishing an empirical relationship between the values of τ and matrix characteristics, we propose a method for the optimal selection of this parameter. Through the generated Figures and the obtained results, it was found that using a rational power of matrix singular values as the τ parameter can lead to an optimal solution for



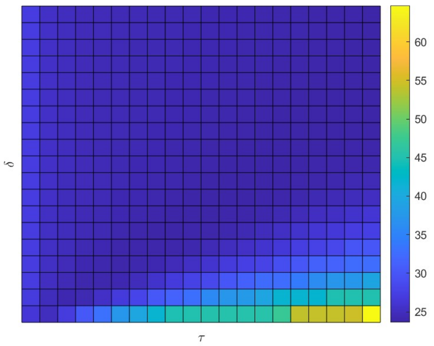
(a) A 128×128 image size



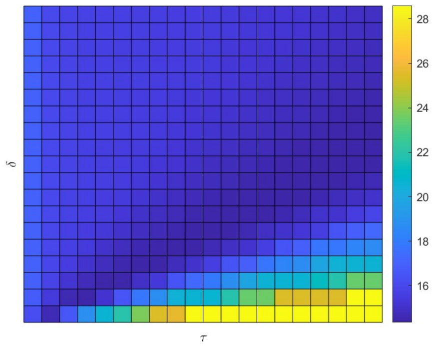
(b) A 256×256 image size



(c) A 512×512 image size

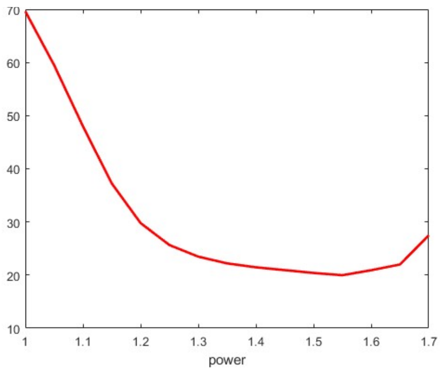


(d) A 200×150 image size

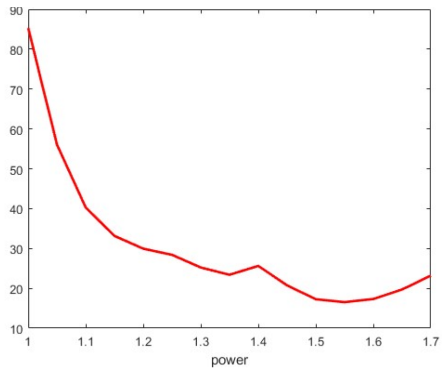


(e) A 180×280 image size

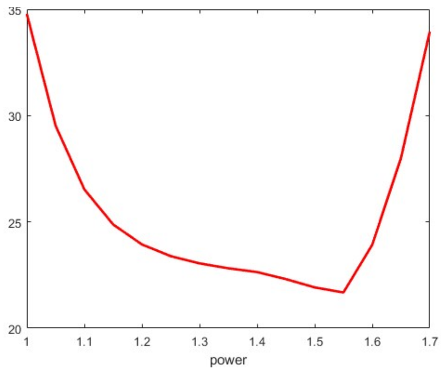
Fig. 5: 2D plots illustrating the variations in error concerning changes in δ and τ for various images at a constant $Sr = 0.5$. The parameter δ ranges from 0.1 to 1.9 with a step size of 0.1, while τ ranges from 10000 to 300000 with a step size of 10000.



(a) $Sr = 0.3$



(b) $Sr = 0.6$



(c) $Sr = 0.9$

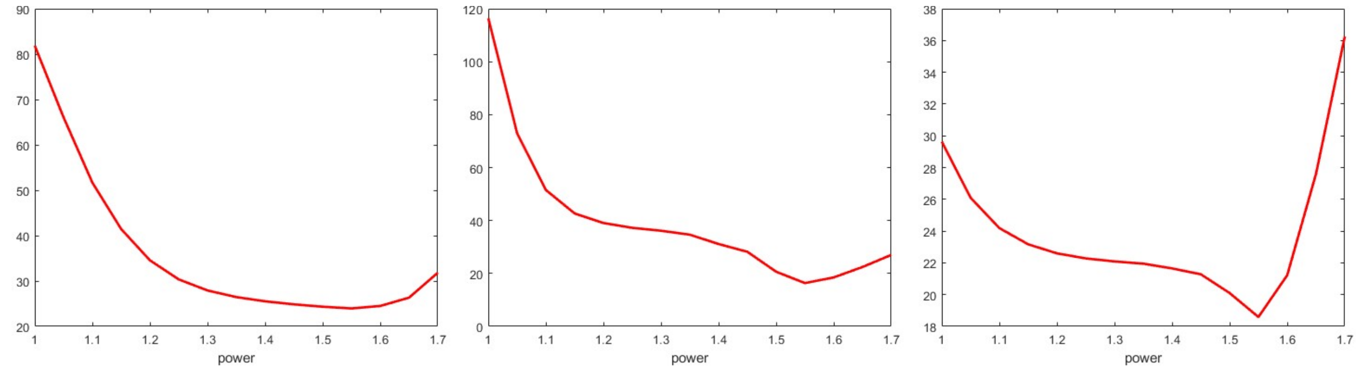
Fig. 6: Plots illustrating the relationship between $power$ and e for a 128×128 image size with varying Sr values. Here, $\tau = \sigma_i^{power}$ and $Sr = i \times 10^{-1}$. The horizontal axis represents $power$, while the vertical axis denotes e.

the matrix completion problem. Inspired by this study, similar ideas can be employed for different problems based on their characteristics and correlations between values to optimize hyperparameters in the model.

5. REFERENCES

[1] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on optimization*, 20(4):1956–1982, 2010.

[2] Emmanuel Candes and Benjamin Recht. Exact matrix com-



(a) $Sr = 0.3$ (b) $Sr = 0.6$ (c) $Sr = 0.9$

Fig. 7: Plots illustrating the relationship between $power$ and e for a 256×256 image size with varying Sr values. Here, $\tau = \sigma_i^{power}$ and $Sr = i \times 10^{-1}$. The horizontal axis represents $power$, while the vertical axis denotes e.

	row	col	Sr	δ^*	τ^*	e^*	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9
row	1.00	0.93	-0.00	-0.15	0.01	0.04	0.56	0.58	0.60	0.64	0.64	0.65	0.65	0.65	0.64
col	0.93	1.00	-0.00	-0.15	0.04	-0.05	0.57	0.57	0.61	0.65	0.66	0.67	0.67	0.67	0.66
Sr	-0.00	-0.00	1.00	-0.24	-0.42	-0.33	0.17	0.14	0.10	0.07	0.06	0.04	0.03	0.03	0.02
δ^*	-0.15	-0.15	-0.24	1.00	0.84	0.03	-0.56	-0.52	-0.48	-0.42	-0.37	-0.33	-0.29	-0.25	-0.23
τ^*	0.01	0.04	-0.42	0.84	1.00	0.06	-0.48	-0.45	-0.39	-0.32	-0.26	-0.21	-0.17	-0.13	-0.10
e^*	0.04	-0.05	-0.33	0.03	0.06	1.00	-0.18	-0.06	-0.14	-0.19	-0.25	-0.28	-0.32	-0.34	-0.37
σ_1	0.56	0.57	0.17	-0.56	-0.48	-0.18	1.00	0.84	0.92	0.90	0.89	0.86	0.84	0.81	0.79
σ_2	0.58	0.57	0.14	-0.52	-0.45	-0.06	0.84	1.00	0.91	0.90	0.83	0.79	0.75	0.71	0.68
σ_3	0.60	0.61	0.10	-0.48	-0.39	-0.14	0.92	0.91	1.00	0.97	0.94	0.91	0.88	0.85	0.82
σ_4	0.64	0.65	0.07	-0.42	-0.32	-0.19	0.90	0.90	0.97	1.00	0.98	0.96	0.94	0.91	0.89
σ_5	0.64	0.66	0.06	-0.37	-0.26	-0.25	0.89	0.83	0.94	0.98	1.00	0.99	0.98	0.96	0.95
σ_6	0.65	0.67	0.04	-0.33	-0.21	-0.28	0.86	0.79	0.91	0.96	0.99	1.00	0.99	0.99	0.97
σ_7	0.65	0.67	0.03	-0.29	-0.17	-0.32	0.84	0.75	0.88	0.94	0.98	0.99	1.00	1.00	0.99
σ_8	0.65	0.67	0.03	-0.25	-0.13	-0.34	0.81	0.71	0.85	0.91	0.96	0.99	1.00	1.00	1.00
σ_9	0.64	0.66	0.02	-0.23	-0.10	-0.37	0.79	0.68	0.82	0.89	0.95	0.97	0.99	1.00	1.00

Fig. 8: The correlation matrix for 50 partial images data with $Sr \in \{0.1, \ldots, 0.9\}$

pletion via convex optimization. *Communications of the ACM*, 55(6):111–119, 2012.

[3] Maryam Fazel. *Matrix rank minimization with applications*. PhD thesis, PhD thesis, Stanford University, 2002.

[4] Massimo Fornasier, Holger Rauhut, and Rachel Ward. Low-rank matrix recovery via iteratively reweighted least squares minimization. *SIAM Journal on Optimization*, 21(4):1614–1640, 2011.

[5] Jincai Ha, Chuan Li, Xiaohu Luo, and Zhi Wang. Matrix completion via modified schatten 2/3-norm. *EURASIP Journal on Advances in Signal Processing*, 2023(1):62, 2023.

[6] Haifeng Wang, Jinchi Chen, and Ke Wei. Implicit regularization and entrywise convergence of riemannian optimization for low tucker-rank tensor completion. *Journal of Machine Learning Research*, 24(347):1–84, 2023.

[7] You-Wei Wen, Kexin Li, and Hefeng Chen. Accelerated matrix completion algorithm using continuation strategy and randomized svd. *Journal of Computational and Applied Mathematics*,

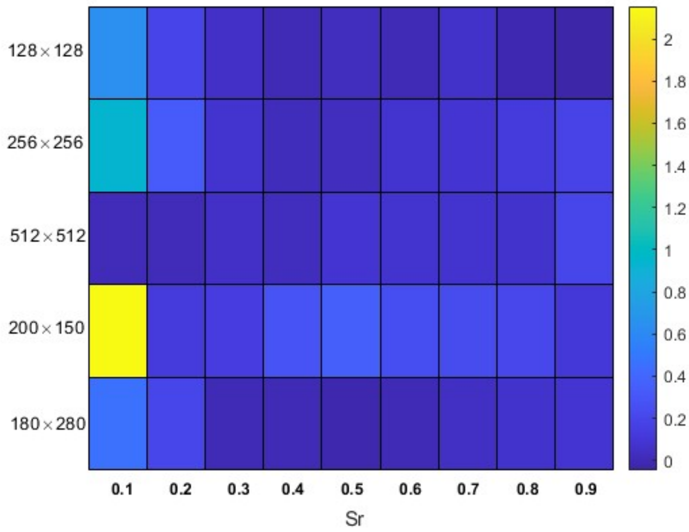


Fig. 9: The mean difference in error between the optimal solution and the solution obtained by setting $\tau = \sigma_i^{1.55}$ for each dataset within various Sr categories ranging from 1 to 9. Rows of Figure correspond to data with different sizes.

$\mathbf{Sr}$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\mathbf{e}_\tau - \mathbf{e}^*$	0.84	0.17	0.06	0.06	0.09	0.08	0.1	0.09	0.1

Table 1. : Comparison of average $e_\tau - e^*$ across 50 distinct data with varying sample rates.

429:115215, 2023.

[8] Quan Yu, Xinzhen Zhang, Yannan Chen, and Liqun Qi. Low tucker rank tensor completion using a symmetric block coordinate descent method. *Numerical Linear Algebra with Applications*, 30(3):e2464, 2023.