Article

Geometric Matrix Completion via Graph-Based Truncated Norm Regularization for Learning Resource Recommendation

Yazhi Yang [1,2,3], Jiandong Shi [4], Siwei Zhou [5] and Shasha Yang [5,*]

1   Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610041, China; yyzhi@mail.cdtu.edu.cn
2   School of Computer Engineering, Chengdu Technological University, Chengdu 611730, China
3   University of Chinese Academy of Sciences, Beijing 100049, China
4   Department of Computer Science, Zhejiang Normal University, Jinhua 321004, China; shijiandong@zjnu.edu.cn
5   Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua 321004, China; siweizhou@zjnu.edu.cn
*   Correspondence: yangss@zjnu.edu.cn

Abstract: In the competitive landscape of online learning, developing robust and effective learning resource recommendation systems is paramount, yet the field faces challenges due to high-dimensional, sparse matrices and intricate user–resource interactions. Our study focuses on geometric matrix completion (GMC) and introduces a novel approach, graph-based truncated norm regularization (GBTNR) for problem solving. GBTNR innovatively incorporates truncated Dirichlet norms for both user and item graphs, enhancing the model's ability to handle complex data structures. This method synergistically combines the benefits of truncated norm regularization with the insightful analysis of user–user and resource–resource graph relationships, leading to a significant improvement in recommendation performance. Our model's unique application of truncated Dirichlet norms distinctively positions it to address the inherent complexities in user and item data structures more effectively than existing methods. By bridging the gap between theoretical robustness and practical applicability, the GBTNR approach offers a substantial leap forward in the field of learning resource recommendations. This advancement is particularly critical in the realm of online education, where understanding and adapting to diverse and intricate user–resource interactions is key to developing truly personalized learning experiences. Moreover, our work includes a thorough theoretical analysis, complete with proofs, to establish the convergence property of the GMC-GBTNR model, thus reinforcing its reliability and effectiveness in practical applications. Empirical validation through extensive experiments on diverse real-world datasets affirms the model's superior performance over existing methods, marking a groundbreaking advancement in personalized education and deepening our understanding of the dynamics in learner–resource interactions.

Keywords: geometric matrix completion; graph-based truncated norm regularization; learning resource recommendation; intelligent education technology and application

MSC: 68T01; 97P80; 15A83

1. Introduction

The digital revolution in education has necessitated the development of intelligent learning resource recommendation systems that can navigate the complex landscape of online learning [1–5]. Central to these systems is the technique of matrix completion, which predicts user preferences by reconstructing a sparsely filled matrix representing user–item interactions [6,7]. While effective, this approach is not without its challenges: the matrices in question are typically high-dimensional and suffer from severe sparsity, which can lead to overfitting and a lack of robustness in the presence of noisy and incomplete data.

Furthermore, conventional matrix completion methods often fail to capture the nuanced relationships and patterns inherent in educational data, which are vital for producing meaningful recommendations that can truly enhance the learning experience [8,9].

From the perspective of real-world application, learning resource recommendation systems often grapple with sparse and high-dimensional data, which can impede the accuracy of recommendations. Traditional recommendation techniques may fall short in effectively leveraging the limited user–item interaction data. To overcome these obstacles, researchers have begun exploring the potential of graph-based methods, which introduce a new dimension to the problem by considering the geometric structure of the data [10–14]. By employing a graph-based framework, geometric matrix completion (GMC) can discern the complex and often subtle relationships within the data, offering a nuanced understanding of user preferences and item characteristics. This graph-centric approach is particularly adept at navigating the sparsity of data, enabling the system to make more informed and accurate predictions, thus significantly enhancing the user experience in educational settings. Technically, this kind of methodology acknowledges that user–item interactions are not merely discrete points within a matrix but are instead part of a rich, interconnected graph that represents the complex relationships between users and learning resources. By integrating this graph structure into the matrix completion process, GMC seeks to leverage the additional information to produce more accurate and relevant recommendations [11,15]. The benefits of this method are manifold, including improved recommendation quality in the face of data sparsity and a more profound understanding of the underlying data geometry. It is within this innovative framework that the next generation of learning resource recommendation systems is being developed, promising a more personalized and insightful approach to online learning.

In our study, we introduce a novel model-driven framework for geometric matrix completion (GMC), with a dedicated focus on the task of learning resource recommendation. Central to our methodology (i.e., GMC-GBTNR) is the introduction of graph-based truncated norm regularization (termed GBTNR; see Definitions 2 and 3 in Section 4.1), a novel technique that redefines the GMC problem and offers a fresh perspective on the underlying geometry of user–learning resource interactions. This regularization framework not only captures the complex topological structures within the data but also adeptly handles the inherent noise and sparsity. We propose a new algorithm that operates within this framework, and we provide a rigorous theoretical analysis of its convergence properties, ensuring that our approach is grounded in solid mathematical principles. Through a series of experiments on five real-world learning resource datasets, we validate the effectiveness of our method. Our findings demonstrate a marked improvement over several established matrix completion techniques, showcasing the practical advantages and theoretical soundness of our proposed model-driven framework. We further substantiate the convergence capability of the GMC-GBTNR model with empirical evidence derived from its performance across five distinct datasets.

In summary, our contributions are as follows:

- We introduce the entirely novel concept of graph-based truncated norm regularization (GBTNR), which lays the foundation for a redefined GMC problem statement, enriching the analytical landscape of recommendation systems.
- In response to the optimization challenges presented by the new GMC formulation, we develop an iterative solution algorithm based on the alternating direction method of multipliers (ADMM), tailored to navigate the complexities of the proposed model.
- We present a thorough theoretical proof that underpins the convergence of our model, ensuring its reliability and efficacy.

The remainder of this paper is organized as follows: Section 2 reviews the related studies, encompassing existing methods for both matrix completion and geometric matrix completion. Section 3 introduces the preliminaries, establishing the necessary background and notations. Section 4 details our proposed method and the corresponding theoretical analysis, presenting the novel graph-based truncated norm regularization (GBTNR)

approach. Section 5 discusses the experiments conducted to validate the efficacy of our approach, including comparisons with established methods. Finally, Section 6 concludes the paper, summarizing our contributions and suggesting avenues for future research.

2. Related Work

This section provides an in-depth review of the existing literature in the areas of matrix completion and its geometric extensions, highlighting seminal studies and recent advancements.

2.1. Matrix Completion

Matrix completion has emerged as a pivotal technique in the realm of machine learning, particularly for its applications in collaborative filtering and recommendation systems [6,7]. The field gained significant momentum following the influential work of Koren et al. [16], who demonstrated its potent capabilities in the Netflix Prize competition. Their approach addressed the inherent challenge of extremely sparse datasets, characteristic of many real-world scenarios, particularly those involving user–item interactions with millions of dimensions. The essence of matrix completion lies in its ability to predict the missing entries of a partially observed matrix with the assumption that the matrix is of low rank. Candès and Recht [17] provide one of the first theoretical guarantees for matrix completion, showing that under certain conditions, it is possible to perfectly reconstruct low-rank matrices from a sample of their entries. This groundbreaking work laid the mathematical foundation for many subsequent algorithms and applications. Subsequent studies have aimed to improve upon the scalability and accuracy of these methods, with some focusing on the use of auxiliary information [18], and others on algorithmic innovations like singular value decomposition (SVD) for enhancing computational efficiency [16]. For example, Hu et al. [19] propose a fast and accurate approach to matrix completion via truncated nuclear norm regularization. Their method balances computational efficiency with the accuracy of reconstruction, contributing to the practical applicability of matrix completion in large-scale problems. Similarly, Li et al. [20] study the error characteristics of matrix elastic-net regularization algorithms, providing valuable insight into the trade-offs involved in regularization parameter selection. Furthering the development of regularization techniques, Liu et al. [21] introduce a truncated nuclear norm regularization method based on weighted residual error, which enhances the matrix completion process. Their work underscores the importance of developing tailored regularization methods to deal with the nuances of matrix completion. In recent years, the realm of matrix completion has witnessed extensions into more intricate formulations. These include challenging variants such as quaternion matrix completion, which proves valuable in addressing large-scale color image and video inpainting tasks [22]. Additionally, there is the exploration of low-rank high-order tensor completion, particularly relevant in handling complex visual data scenarios [23]. An alternative direction delves into inductive matrix completion, with promising applications in the domain of wireless communications [24].

2.2. Geometric Matrix Completion

Recent strides in matrix completion have been influenced by the integration of geometric structures, which offer a nuanced perspective on user–item interactions. Ma et al. [25] are among the first to suggest the use of manifold learning techniques to exploit the geometric relationships within the data. By borrowing ideas from the field of manifold learning [26,27], Kalofolias et al. [10] extend this by utilizing graph-based approaches to encode similarities, thereby harnessing the inherent structure of the data for improved matrix completion. Rao et al. [28] and Kuang et al. [29] further explored the potential of graph regularization, underscoring the importance of preserving local structures and smoothness in the recommendation process. These contributions are deeply rooted in the field of signal processing on graphs, where Shuman et al. [30] have provided extensive insights into extending classical signal processing techniques to graph-structured data.

The concept of smoothness on graphs has been particularly influential, suggesting that similar users and items (learning resources) are likely to have similar preferences or characteristics. This notion has been operationalized through the development of graph Laplacian regularization techniques, which have been shown to enhance the performance of recommender systems by promoting smoothness over the graph [31,32]. Such techniques have also been instrumental in the formulation of graph convolutional networks [33], which have successfully adapted deep learning methods to data represented on irregular domains. Furthermore, certain research endeavors delve into the matrix completion problem while considering the underlying geometric or topological relationships among rows and/or columns. This is approached from various angles, including spectral viewpoints as explored in the work in [34], a Riemannian manifold perspective as investigated in [35], and the utilization of Plücker coordinates as highlighted in [36].

Summary. Technically, our work represents a distinctive stride in the landscape of matrix completion research, particularly in the realm of geometric matrix completion. We introduce a novel approach, graph-based truncated norm regularization (GBTNR), which combines truncated nuclear norm regularization with graph information. This integration is specifically tailored to address the challenges of learning resource recommendation. GBTNR is designed to adeptly mine and utilize the latent relationships not only between users but also among learning resources themselves. Our methodology transcends traditional limitations of matrix completion techniques, such as their vulnerability to noise and propensity to overestimate the matrix rank. By leveraging the strengths of both graph-based structures and truncated norm regularization, we present a solution that is not only robust and theoretically sound but also uniquely suited to enhance recommendation systems in the e-learning sector. This work, therefore, stands as a significant contribution to the ongoing evolution of matrix completion strategies, particularly in their application to educational technologies.

3. Preliminaries

This section lays the groundwork for understanding the intricacies of matrix completion and geometric matrix completion problems. We delve into the theoretical and methodological preliminaries essential for addressing the challenge of deducing the unknown entries of a matrix from a sparse subset of its observed elements. The ensuing discussion bifurcates into two principal domains: the classical matrix completion framework, which harnesses the power of low-rank matrix assumptions, and the more nuanced geometric matrix completion approach, which incorporates the geometry of the data into the recovery process.

Matrix completion problem. The challenge of inferring the missing elements of a matrix from a subset of its entries necessitates additional mathematical stipulations to define a solvable problem space. A common approach is to posit that the data are restricted to a lower-dimensional subspace, indicative of the matrix's low rank:

$$\min_{X} \ \text{rank}(X) \quad \text{s.t.} \quad x_{ij} = y_{ij}, \ \forall ij \in \Omega, \tag{1}$$

where X represents the matrix under reconstruction, Ω encapsulates the indices of known values, and y_{ij} are the corresponding known values. To accommodate noise and variations, the strict equality can be substituted with a penalization term:

$$\min_{X} \ \text{rank}(X) + \frac{\mu}{2}\|\Omega \circ (X - Y)\|_{\text{F}}^2, \tag{2}$$

where Ω denotes the indicator matrix, and $\circ$ signifies the Hadamard product.

Yet, minimizing the rank is an NP-hard problem, rendering it unsolvable for large-scale applications. The most effective convex surrogate for this is:

$$\min_{X} \ \|X\|_* + \frac{\mu}{2}\|\Omega \circ (X - Y)\|_{\text{F}}^2, \tag{3}$$

with $\|\cdot\|_*$ denoting the nuclear norm, summing up the singular values of the matrix, as established in the seminal study [17]. Theoretically and empirically, Candès and Recht [17] demonstrate that employing the ℓ_1 norm as a relaxation technique in singular value decomposition significantly enhances the likelihood of accurately reconstructing the original matrix with low-rank characteristics.

Geometric matrix completion. An alternative formulation applies a geometric perspective to the rank minimization (see Equation (1)), constraining the solution space to adhere to an inherent geometric structure among the matrix's rows and columns [10,25,28,37]. The fundamental model is based on the concept of proximity, characterized as an undirected weighted graph for columns, denoted as $\mathcal{G}_c = (1, \ldots, n, \mathcal{E}_c, W_c)$, which is defined by an adjacency matrix W_c. The elements of this matrix, w_{ij}^c, are symmetric and only non-zero if an edge exists between nodes i and j within the edge set $\mathcal{E}_c$, representing a connection in the graph. Similarly, the row graph $\mathcal{G}_r$, which encapsulates item similarities, is constructed in a like manner and can be represented by $\mathcal{G}_r = (1, \ldots, m, \mathcal{E}_r, W_r)$. In the context of recommending learning resources, the row graph can be perceived as analogous to a social network where the edges mirror the relationships between individual students and the affinities they share. Conversely, the column graph represents the intricate web of connections between various learning resources, capturing the similarities or thematic overlaps between them. This dual-graph structure can provide useful auxiliary information for recommendation algorithms, where student–student ties potentially reflect common learning interests and resource–resource links represent curricular similarities.

For both graphs, one can formulate the unnormalized graph Laplacian, a symmetric positive-semidefinite matrix $\Delta = I - D^{-1/2}WD^{-1/2}$, with D being the degree matrix, calculated as $\text{diag}(\sum_{j \neq i} w_{ij})$. The Laplacians pertinent to the row and column graphs are symbolized as Δ_r and Δ_c, respectively. This allows us to measure the smoothness of the matrix's columns or rows seen as functions on the respective graphs. This smoothness is quantified by the Dirichlet form: $\|X\|_{\mathcal{G}_r}^2 = \text{trace}(X^\top \Delta_r X)$ (respecitvely, $\|X\|_{\mathcal{G}_c}^2 = \text{trace}(X\Delta_c X^\top)$).

The geometric matrix completion aims to minimize:

$$\min_X \ \|X\|_{\mathcal{G}_r}^2 + \|X\|_{\mathcal{G}_c}^2 + \frac{\mu}{2}\|\Omega \circ (X - Y)\|_{\text{F}}^2, \tag{4}$$

Definition 1. *The singular value decomposition (SVD) of a matrix $X \in \mathbb{R}^{m \times n}$ takes the form (see [38,39]):*

$$X = U\Sigma V^T, \Sigma = \text{diag}\left(\{\sigma_i\}_{1 \leq i \leq \min(m,n)}\right),$$

where U and V have orthonormal columns, and Σ is diagonal with singular values. For $\alpha > 0$, the singular value shrinkage operator S_α is defined via:

$$S_\alpha(X) = US_\alpha(\Sigma)V^T,$$

where $S_\alpha(\Sigma) = \text{diag}\left(\{\max(\sigma_i - \alpha, 0)\}_{1 \leq i \leq \min(m,n)}\right)$.

Proposition 1. *Given any $\alpha > 0$ and matrix $Q \in \mathbb{R}^{m \times n}$, the singular value shrinkage operator satisfies (see [38,39]):*

$$S_\alpha(Q) = \arg\min_X \alpha\|X\|_* + \frac{1}{2}\|X - Q\|_{\text{F}}^2.$$

4. Geometric Matrix Completion via Hybrid Truncated Norm Regularization

4.1. Proposed Method

Basics. Given a matrix $X \in \mathbb{R}^{m \times n}$, we assume that its geometric information, that is, the row graph $\mathcal{G}_r$ and the row graph $\mathcal{G}_c$, are available. Let $A_r \in \mathbb{R}^{m \times m}$ denote the adjacency matrix of $\mathcal{G}_r$ and $A_c \in \mathbb{R}^{n \times n}$ denote the adjacency matrix of $\mathcal{G}_c$. Their (normalized) graph Laplacian matrices are expressed as $\mathcal{L}_r := I - D_r^{-1/2}A_r D_r^{-1/2}$ and $\mathcal{L}_c := I - D_c^{-1/2}A_c D_c^{-1/2}$, respectively, where D_r (and D_c) is the degree matrix of $\mathcal{G}_r$ (and

$\mathcal{G}_c$). Based on spectral graph theory [40], we can have their Laplacian eigendecompositions: $\mathcal{L}_r = \Phi_r \Lambda_r \Phi_r{}^T$ and $\mathcal{L}_c = \Phi_c \Lambda_c \Phi_c{}^T$, respectively, where Φ. denotes the matrix of orthonormal eigenvectors and Λ. is the diagonal matrix of the corresponding eigenvalues. $\Lambda_r{}^{1/2}$ (and $\Lambda_c{}^{1/2}$) is well-defined due to the positive-semidefiniteness of $\mathcal{L}_r$ (and $\mathcal{L}_c$).

New Definitions. With these notations, we define the truncated Dirichlet norms for the row and column graphs as follows.

Definition 2. *The truncated Dirichlet norm for the row graph $\mathcal{G}_r$ is defined as*

$$\|X\|_{\mathcal{G}_r,\eta} = \sqrt{\sum_{i=\eta+1}^{\min(m,n)} \sigma_i^2 \left(\Lambda_r{}^{1/2} \Phi_r{}^T X \right)}. \tag{5}$$

Definition 3. *The truncated Dirichlet norm for the column graph $\mathcal{G}_c$ is defined as*

$$\|X\|_{\mathcal{G}_c,\eta} = \sqrt{\sum_{i=\eta+1}^{\min(m,n)} \sigma_i^2 \left(X \Phi_c \Lambda_c{}^{1/2} \right)}. \tag{6}$$

Our proposed GMC-GBTNR. Our proposed method can be formulated as:

$$\min_X \|X\|_{\eta,*} + \lambda_1 \|X\|_{\mathcal{G}_r,\eta}^2 + \lambda_2 \|X\|_{\mathcal{G}_c,\eta}^2$$
$$\text{s. t. } P_\Omega(X) = P_\Omega(M), \tag{7}$$

where $\lambda_1 > 0$, $\lambda_2 > 0$ are regularization parameters.

As shown in Equation (7), our GBTNR approach leverages truncated Dirichlet norms as regularization terms in the optimization problem for matrix completion. Specifically, these norms act on the row and column graphs, incorporating the geometric structure of the data into the learning process. They contribute to the regularization by emphasizing only the significant singular values, as represented by the sum of the squared singular values multiplied by their corresponding truncated eigenvectors, starting from the $(\eta + 1)$-th smallest one, as per the truncation parameter η.

Theoretically, the truncated Dirichlet norms in the GBTNR approach encapsulate the graph information of users and items, enhancing the model performance. These terms, acting as regularization in the optimization framework, exploit the relational structure inherent in the user–item interactions. This is particularly beneficial in situations with sparse rating data, where traditional collaborative filtering might struggle. By integrating user and item graph information, the GBTNR method can discern underlying patterns and relationships, which are crucial for accurate matrix completion. The technique effectively harnesses these relational data, often leading to more precise recommendations.

Proposition 2. *For any given matrix $X \in R^{m \times n}$, any matrices $A \in R^{\eta \times m}$, $B \in R^{\eta \times n}$ satisfying $AA^T = I$ (I is the unit matrix of $\eta \times \eta$ and $\eta \leq \min(m,n)$), we have (see [19])*

$$Tr(AXB^T) \leq \sum_{i=1}^{\eta} \sigma_i(X). \tag{8}$$

Proposition 3. *For any given matrix $X \in R^{m \times n}$, any matrix $A \in R^{\eta \times m}$ satisfying $AA^T = I$ (I is the unit matrix of $\eta \times \eta$ and $\eta \leq \min(m,n)$), we have (see [41])*

$$\|AX\|_F^2 \leq \sum_{i=1}^{\eta} \sigma_i^2(X). \tag{9}$$

For simplicity, we denote $\mathcal{T}_r := \Lambda_r^{1/2}\Phi_r^T$ as a transform matrix induced by row graph $\mathcal{G}_r$ and $\mathcal{T}_c := \Phi_c\Lambda_c^{1/2}$ as a transform matrix induced by the column graph $\mathcal{G}_c$. Suppose $\tilde{U}\tilde{\Sigma}\tilde{V}^T$ is the SVD of $\mathcal{T}_r X$ and $\hat{U}\hat{\Sigma}\hat{V}^T$ is the SVD of $X\mathcal{T}_c$ with $\tilde{U} = [\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_m] \in R^{m\times m}$, $\hat{U} = [\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_m] \in R^{m\times m}$. Given the SVD of X, i.e, $U\Sigma V^T$, where $U = [u_1, u_2, \cdots, u_m] \in R^{m\times m}, \Sigma \in R^{m\times n}$, and $V = [v_1, v_2, \cdots, v_n] \in R^{n\times n}$, we can take

$$A = [u_1, u_2, \cdots, u_\eta]^T, \quad B = [v_1, v_2, \cdots, v_\eta]^T,$$
$$C = [\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_\eta]^T, \quad D = [\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_\eta]^T. \tag{10}$$

Based on Propositions 2 and 3, it is easy to verify that

$$\max_{\substack{AA^T=I\\BB^T=I}} \text{Tr}(AXB^T) = \sum_{i=1}^{\eta} \sigma_i(X),$$

$$\max_{CC^T=I} \|C\mathcal{T}_r X\|_F^2 = \sum_{i=1}^{\eta} \sigma_i^2(\mathcal{T}_r X),$$

$$\max_{DD^T=I} \|DX\mathcal{T}_c\|_F^2 = \sum_{i=1}^{\eta} \sigma_i^2(X\mathcal{T}_c)$$

Then the objective function (7) can be reformulated as:

$$\min_X \|X\|_* - \max_{\substack{AA^T=I\\BB^T=I}} \text{Tr}(AXB^T)$$
$$+ \lambda_1 (\|\mathcal{T}_r X\|_F^2 - \max_{CC^T=I} \|C\mathcal{T}_r X\|_F^2)$$
$$+ \lambda_2 (\|X\mathcal{T}_c\|_F^2 - \max_{DD^T=I} \|DX\mathcal{T}_c\|_F^2) \tag{11}$$
$$\text{s. t. } P_\Omega(X) = P_\Omega(M).$$

Using the alternative optimization scheme used in [19,42], we can solve the optimization problem (11) via two steps. First, we initialize $X_1 = M_\Omega$. In the l-th iteration, we fix X_l and calculate two intermediate matrices $\mathcal{T}_r X_l$ and $X_l\mathcal{T}_c$, then conduct SVDs of X_l, $\mathcal{T}_r X$, and $X\mathcal{T}_c$, respectively, to obtain A_l, B_l, C_l, D_l as formulated in (10). Second, we fix A_l, B_l, C_l, and D_l to update X_{l+1} based on the following problem:

$$\min_X \|X\|_* - \text{Tr}(A_l X B_l^T) + \lambda_1 (\|\mathcal{T}_r X\|_F^2 - \|C_l\mathcal{T}_r X\|_F^2)$$
$$+ \lambda_2 (\|X\mathcal{T}_c\|_F^2 - \|D_l X\mathcal{T}_c\|_F^2) \tag{12}$$
$$\text{s. t. } P_\Omega(X) = P_\Omega(M).$$

Algorithm 1 presents the detailed pseudo-code for these two steps. It should be noted that we assume $\mathcal{T}_r$, $\mathcal{T}_c$ are given as inputs for the algorithm implementation, provided that both the row graph $\mathcal{G}_r$ and the column graph $\mathcal{G}_c$ are known in advance. In practice, it is possible to only consider one transform matrix, i.e., $\mathcal{T}_r$ or $\mathcal{T}_c$ for the case where only $\mathcal{G}_r$ or $\mathcal{G}_c$ is available by setting $\lambda_2 = 0$ or $\lambda_1 = 0$. The extreme case in which both $\mathcal{T}_r$ and $\mathcal{T}_c$ are not provided does not fit our problem formulation, which relies on the basis of geometric matrix completion. In the next section, will discuss how to solve the optimization problem (12).

Algorithm 1 GMC-GBTNR: Geometric Matrix Completion via Graph-Based Truncated Norm Regularization

---

Require: Original incomplete data M_Ω, $\mathcal{T}_r$, $\mathcal{T}_c$, regularization parameters $\lambda_1 > 0$ and $\quad \lambda_2 > 0$, tolerance $\varepsilon_0 > 0$;
Ensure: The recovered matrix X^*;
$\quad$ **Initialize:** $X_1 = P_\Omega(M) = M_\Omega$;
$\quad$ **repeat**
$\quad\quad$ **Step 1:** Given X_l, calculate $\mathcal{T}_r X_l$ and $X_l \mathcal{T}_c$, then
$\quad\quad$ calculate the SVDs of X_l, $\mathcal{T}_r X_l$, and $X_l \mathcal{T}_c$, respectively:

$$[U_l, \Sigma_l, V_l] = \text{SVD}(X_l),$$

$$[\tilde{U}_l, \check{\Sigma}_l, \tilde{V}_l] = \text{SVD}(\mathcal{T}_r X_l),$$

$$[\hat{U}_l, \hat{\Sigma}_l, \hat{V}_l] = \text{SVD}(X_l \mathcal{T}_c),$$

$\quad\quad$ where

$$U_l = [u_1, u_2, \cdots, u_m], \ \ V_l = [v_1, v_2, \cdots, v_m],$$

$$\tilde{U} = [\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_m], \ \ \hat{U} = [\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_m].$$

$\quad\quad$ Take A_l, B_l, C_l, D_l as
$\quad\quad A_l = [u_1, u_2, \cdots, u_\eta]^T, \ \ B_l = [v_1, v_2, \cdots, v_\eta]^T$;
$\quad\quad C_l = [\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_\eta]^T, \ \ D_l = [\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_\eta]^T.$
$\quad\quad$ **Step 2:** Solve the following optimization problem:

$$X_{l+1} = \arg\min_X \|X\|_* - \text{Tr}(A_l X B_l^T) + \lambda_1 (\|\mathcal{T}_r X\|_F^2$$
$$- \|C_l \mathcal{T}_r X\|_F^2) + \lambda_2 (\|X \mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2)$$
$$\text{s. t. } P_\Omega(X_l) = P_\Omega(M).$$

$\quad$ **until** $\|X_{l+1} - X_l\|_F \leq \varepsilon_0$.
$\quad$ **Return** $X^* = X_{l+1}$

4.2. Optimization Algorithms

In this subsection, an efficient iterative optimization scheme is created to optimize (12). According to the augmented Lagrangian multiplier method, a common strategy is to approximately minimize the augmented Lagrangian function by adopting an alternating scheme and we use it to solve the model (12). To this end, we introduce the auxiliary variable Z and then (12) is equivalent to the following form:

$$\min_X \|X\|_* - \text{Tr}(A_l W B_l^T) + \lambda_1 (\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2)$$
$$+ \lambda_2 (\|X \mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) \tag{13}$$
$$\text{s. t. } X = W, P_\Omega(W) = P_\Omega(M).$$

Note that $X = W$ and $P_\Omega(W) = P_\Omega(M)$ are two linear constraints. To handle the two constraints simultaneously, we reformulate the model (13) as follows:

$$\min_X \|X\|_* - \text{Tr}(A_l W B_l^T) + \lambda_1 (\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2)$$
$$+ \lambda_2 (\|X \mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) \tag{14}$$
$$\text{s. t. } \mathcal{A}X + \mathcal{B}W = Z,$$

where $\mathcal{A}, \mathcal{B} : R^{m \times n} \to R^{2m \times 2n}$ are two linear operators denoted by

$$\mathcal{A}X = \begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix}, \mathcal{B}W = \begin{bmatrix} -W & 0 \\ 0 & P_\Omega(W) \end{bmatrix}, \tag{15}$$

and

$$Z = \begin{bmatrix} 0 & 0 \\ 0 & P_\Omega(M) \end{bmatrix}, \tag{16}$$

respectively.

Given a matrix Y with the following block representation,

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}, \tag{17}$$

where $Y_{ij} \in R^{m \times n}$ ($i, j = 1, 2$), it is easy to verify that the adjoint operators of $\mathcal{A}$ and $\mathcal{B}$, that is, $\mathcal{A}^*$ and $\mathcal{B}^*$, respectively, can be evaluated by (see [19,42])

$$\mathcal{A}^* Y = Y_{11}, \quad \mathcal{B}^* Y = -Y_{11} + P_\Omega(Y_{22}). \tag{18}$$

Then, the augmented Lagrange function of (14) can be written as:

$$
\begin{aligned}
\pounds_\mu(X, W, Y) =& \|X\|_* - \mathrm{Tr}(A_l W B_l^T) + \lambda_1(\|\mathcal{T}_r X\|_F^2 \\
&- \|C_l \mathcal{T}_r X\|_F^2) + \lambda_2(\|X\mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) \\
&+ \langle Y, \mathcal{A}X + \mathcal{B}W - Z \rangle + \frac{\mu}{2}\|\mathcal{A}X + \mathcal{B}W - Z\|_F^2,
\end{aligned}
\tag{19}
$$

where Y is the Lagrange multiplier matrix and $\mu > 0$ is the penalty parameter. Thus, we can adopt the alternating iteration strategy, i.e., fix some variables and solve the remaining one.

Initially, we compute X^{k+1}. We can give the following solution by ignoring the constant term:

$$
\begin{aligned}
X^{(k+1)} =& \arg\min_X \pounds_\mu(X, W^{(k)}, Y^{(k)}) \\
=& \arg\min_X \|X\|_* - \mathrm{Tr}(A_l W B_l^T) + \lambda_1(\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2) \\
&+ \lambda_2(\|X\mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) + \langle Y^{(k)}, \mathcal{A}X + \mathcal{B}W^{(k)} - C \rangle \\
&+ \frac{\mu}{2}\|\mathcal{A}X + \mathcal{B}W^{(k)} - Z\|_F^2 \\
=& \arg\min_X \|X\|_* + \lambda_1(\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2) + \lambda_2(\|X\mathcal{T}_c\|_F^2 \\
&- \|D_l X \mathcal{T}_c\|_F^2) + \frac{\mu}{2}\|\mathcal{A}X + \mathcal{B}W^{(k)} - Z + \frac{Y^{(k)}}{\mu}\|_F^2.
\end{aligned}
\tag{20}
$$

Denote a differentiable real function with respect to its matrix argument X as follows:

$$
\begin{aligned}
\mathcal{F}(X) =& \lambda_1(\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2) + \lambda_2(\|X\mathcal{T}_c\|_F^2 \\
&- \|D_l X \mathcal{T}_c\|_F^2) + \frac{\mu}{2}\|\mathcal{A}X + \mathcal{B}W^{(k)} - Z + \frac{Y^{(k)}}{\mu}\|_F^2.
\end{aligned}
$$

Then, $\mathcal{F}(X)$ can be approximated by the Taylor series

$$\mathcal{F}(X) = \mathcal{F}(X^{(k)}) + \nabla\mathcal{F}(X^{(k)})(X - X^{(k)}) + \frac{\nu}{2}\|X - X^{(k)}\|_F^2,$$

where $\nu > \mu$ and the first-order gradient $\nabla \mathcal{F}(X^{(k)})$ can be evaluated by

$$
\begin{aligned}
& \nabla \mathcal{F}(X^{(k)}) \\
= \; & \lambda_1 (2\mathcal{T}_r^T \mathcal{T}_r X^{(k)} - 2\mathcal{T}_r^T C_l^T C_l \mathcal{T}_r X^{(k)}) \\
& + \lambda_2 (2X^{(k)} \mathcal{T}_c \mathcal{T}_c^T - 2D_l D_l^T X^{(k)} \mathcal{T}_c \mathcal{T}_c^T) \\
& + Y_{11}^{(k)} - \mu W^{(k)} + \mu X^{(k)} \\
= \; & (2\lambda_1 \mathcal{T}_r^T \mathcal{T}_r - 2\lambda_1 \mathcal{T}_r^T C_l^T C_l \mathcal{T}_r + \mu I) X^{(k)} + \\
& (2\lambda_2 I - 2\lambda_2 D_l D_l^T) X^{(k)} \mathcal{T}_c \mathcal{T}_c^T + Y_{11}^{(k)} - \mu W^{(k)}.
\end{aligned}
\tag{21}
$$

Therefore, the solution (20) can be refined as:

$$
\begin{aligned}
& X^{(k+1)} \\
& = \arg \min_X \|X\|_* + \nabla \mathcal{F}(X^{(k)})(X - X^{(k)}) + \frac{\nu}{2}\|X - X^{(k)}\|_F^2 \\
& = \arg \min_X \|X\|_* + \frac{\nu}{2}\|X - X^{(k)} + \frac{\nabla \mathcal{F}(X^{(k)})}{\nu}\|_F^2.
\end{aligned}
$$

By Definition 1 and Proposition 1, this could be solved as

$$
X^{(k+1)} = S_{\frac{1}{\nu}}\left(X^{(k)} - \frac{\nabla \mathcal{F}(X^{(k)})}{\nu}\right).
\tag{22}
$$

Next, we concentrate on calculating $W^{(k+1)}$.

$$
\begin{aligned}
W^{(k+1)} & = \arg \min_W \pounds_\mu(X^{(k+1)}, W, Y^{(k)}) \\
& = \arg \min_W \frac{\mu}{2}\|\mathcal{A}X^{(k+1)} + \mathcal{B}W - Z\|_F^2 - \mathrm{Tr}(A_l W B_l^T) \\
& \quad + \langle Y^{(k)}, \mathcal{A}X^{(k+1)} + \mathcal{B}W - Z \rangle.
\end{aligned}
\tag{23}
$$

It is obvious that the objective function of (23) is a quadratic function.

$$
\begin{aligned}
W^{(k+1)} = \; & \frac{1}{\mu}\left(A_l^T B_l + Y_{11}^{(k)}\right) + X^{(k+1)} \\
& - \frac{1}{2\mu} P_\Omega\left(\mu(X^{(k+1)} - M) + A_l^T B_l + Y_{11}^{(k)} + Y_{22}^{(k)}\right).
\end{aligned}
\tag{24}
$$

Ultimately, the update of Lagrange multipliers is

$$
Y^{(k+1)} = Y^{(k)} + \mu(\mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z).
\tag{25}
$$

To sum up, the full procedure is listed in Algorithm 2.

Algorithm 2 ADMM for **Step 2** of **Algorithm 1**

Require: A_l, B_l, C_l, D_l, M_Ω and tolerance $\varepsilon > 0$;
    Initialize: $X^{(1)} = M_\Omega$, $W^{(1)} = X^{(1)}$, $Y^{(1)} = X^{(1)}$, μ, and $\nu = 1.1\mu$;
    repeat
      Step 1: Fix $W^{(k)}$ and $Y^{(k)}$, then compute $\mathcal{F}(X^{(k)})$ by (21) and $X^{(k+1)}$ by (22)
      Step 2: Fix $Y^{(k)}$ and $X^{(k+1)}$, then compute $W^{(k+1)}$ by (24)
      Step 3: Fix $W^{(k+1)}$ and $X^{(k+1)}$, then compute $Y^{(k+1)}$ by (25)
    until $\|X^{(k+1)} - X^{(k)}\|_F \le \varepsilon$.
Ensure: $X^{(k+1)}$

4.3. Convergence Analysis

In this subsection, the convergence of the proposed approach will be discussed. To this end, denote the optimal value of (14) as

$$
\begin{aligned}
g^* = \inf\{ & \|X\|_* - \mathrm{Tr}(A_l W B_l^T) + \lambda_1(\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2) \\
& + \lambda_2(\|X\mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) : \mathcal{A}X + \mathcal{B}W = Z\},
\end{aligned}
\tag{26}
$$

and the unaugmented Lagrangian function as

$$
\begin{aligned}
£_0(X, W, Y) = & \|X\|_* - \mathrm{Tr}(A_l W B_l^T) + \lambda_1(\|\mathcal{T}_r X\|_F^2 \\
& - \|C_l \mathcal{T}_r X\|_F^2) + \lambda_2(\|X\mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2).
\end{aligned}
\tag{27}
$$

Subsequently, the paper establishes three auxiliary lemmas as foundational results prior to demonstrating the convergence of the proposed method.

Lemma 1. *Assume that (X^*, W^*, Y^*) is the optimal solution to the unaugmented Lagrangian function (27). Let $\left(X^{(k+1)}, W^{(k+1)}\right)$ be the iterative solution to Algorithm 2 and*

$$
\begin{aligned}
g^{(k+1)} = & \|X^{(k+1)}\|_* - \mathrm{Tr}(A_l W^{(k+1)} B_l^\top) + \lambda_1\left(\|\mathcal{T}_r X^{(k+1)}\|_F^2 - \|C_l \mathcal{T}_r X^{(k+1)}\|_F^2\right) \\
& + \lambda_2\left(\|X^{(k+1)} \mathcal{T}_c\|_F^2 - \|D_l X^{(k+1)} \mathcal{T}_c\|_F^2\right).
\end{aligned}
$$

Then the following inequality holds:

$$
g^* - g^{(k+1)} \le \langle Y^*, \mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z \rangle.
\tag{28}
$$

Proof. Since (X^*, W^*, Y^*) is the optimal solution to (27), then it follows that

$$
£_0(X^*, W^*, Y^*) \le £_0(X^{(k+1)}, W^{(k+1)}, Y^*).
$$

Using $\mathcal{A}X^* + \mathcal{B}W^* = Z$, the left-hand side is equal to g^*. In terms of

$$
\begin{aligned}
g^{(k+1)} = & \|X^{(k+1)}\|_* - \mathrm{Tr}(A_l W^{(k+1)} B_l^\top) + \lambda_1\left(\|\mathcal{T}_r X^{(k+1)}\|_F^2 - \|C_l \mathcal{T}_r X^{(k+1)}\|_F^2\right) \\
& + \lambda_2\left(\|X^{(k+1)} \mathcal{T}_c\|_F^2 - \|D_l X^{(k+1)} \mathcal{T}_c\|_F^2\right),
\end{aligned}
$$

we can obtain that $g^* \le g^{(k+1)} + \langle Y^*, \mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z \rangle$, i.e., the inequality (28) holds. This completes the proof. □

Actually, $\mathcal{A}X^* + \mathcal{B}W^* = Z$ implies that (X^*, W^*) is also the optimal solution to the model (14).

Lemma 2. *Let $E^{(k+1)} = \mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z$ and $Y^{(k+1)} = Y^{(k)} + \mu E^{(k+1)}$. Then we have*

$$
\begin{aligned}
g^{(k+1)} - g^* \le & -\langle Y^{(k+1)}, E^{(k+1)} \rangle - \mu \langle \mathcal{B}(W^{(k+1)} - W^{(k)}), \\
& \mathcal{B}(W^{(k+1)} - W^*) - E^{(k+1)} \rangle.
\end{aligned}
\tag{29}
$$

Proof. By definition, we know that $X^{(k+1)}$ minimizes $£_\mu(X, W^{(k)}, Y^{(k)})$. As $£_\mu(X, W^{(k)}, Y^{(k)})$ is closed, convex, and subdifferentiable on X. With the property of subdifferential [43], the optimality condition can be satisfied, namely,

$$
\begin{aligned}
0 \in \partial \pounds_\mu(X^{(k+1)}, W^{(k)}, Y^{(k)}) \\
= \partial \|X^{(k+1)}\|_* + 2\lambda_1(\mathcal{T}_r^T \mathcal{T}_r X^{(k+1)} - \mathcal{T}_r^T C_l^T C_l \mathcal{T}_r X^{(k+1)}) \\
+ 2\lambda_2(X^{(k+1)} \mathcal{T}_c \mathcal{T}_c^T - D_l D_l^T X^{(k+1)} \mathcal{T}_c \mathcal{T}_c^T) \\
+ \mathcal{A}^* Y^{(k)} + \mu \mathcal{A}^*(\mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k)} - Z).
\end{aligned}
\tag{30}
$$

Using $Y^{k+1} = Y^k + \mu R^{k+1}$, we can rearrange (30) as

$$
\begin{aligned}
0 \in \partial \pounds_\mu(X^{(k+1)}, W^{(k)}, Y^{(k)}) \\
= \partial \|X^{(k+1)}\|_* + 2\lambda_1(\mathcal{T}_r^T \mathcal{T}_r X^{(k+1)} - \mathcal{T}_r^T C_l^T C_l \mathcal{T}_r X^{(k+1)}) \\
+ 2\lambda_2(X^{(k+1)} \mathcal{T}_c \mathcal{T}_c^T - D_l D_l^T X^{(k+1)} \mathcal{T}_c \mathcal{T}_c^T) \\
+ \mathcal{A}^*(Y^{(k+1)} - \mu \mathcal{B}(W^{(k+1)} - W^{(k)})).
\end{aligned}
$$

This indicates that $X^{(k+1)}$ minimizes

$$
\begin{aligned}
\|X\|_* + \lambda_1(\|\mathcal{T}_r X\|_F^2 - \|C_l \mathcal{T}_r X\|_F^2) \\
+ \lambda_2(\|X\mathcal{T}_c\|_F^2 - \|D_l X \mathcal{T}_c\|_F^2) \\
+ \langle Y^{(k+1)} - \mu \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{A}X \rangle.
\end{aligned}
$$

Similarly, $W^{(k+1)}$ minimizes

$$
-\text{Tr}(A_l W B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W \rangle.
$$

Accordingly, for the optimal solution (X^*, W^*), we can obtain

$$
\begin{aligned}
\|X^{(k+1)}\|_* + \lambda_1(\|\mathcal{T}_r X^{(k+1)}\|_F^2 - \|C_l \mathcal{T}_r X^{(k+1)}\|_F^2) \\
+ \lambda_2(\|X^{(k+1)} \mathcal{T}_c\|_F^2 - \|D_l X^{(k+1)} \mathcal{T}_c\|_F^2) \\
+ \langle Y^{(k+1)} - \mu \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{A}X^{(k+1)} \rangle \\
\leq \|X^*\|_* + \lambda_1(\|\mathcal{T}_r X^*\|_F^2 - \|C_l \mathcal{T}_r X^*\|_F^2) \\
+ \lambda_2(\|X^* \mathcal{T}_c\|_F^2 - \|D_l X^* \mathcal{T}_c\|_F^2) \\
+ \langle Y^{(k+1)} - \mu \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{A}X^* \rangle,
\end{aligned}
\tag{31}
$$

and

$$
\begin{aligned}
-\text{Tr}(A_l W^{(k+1)} B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W^{(k+1)} \rangle \\
\leq -\text{Tr}(A_l W^* B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W^* \rangle.
\end{aligned}
\tag{32}
$$

Adding both sides of (31) and (32), as well as using $\mathcal{A}X^* + \mathcal{B}W^* = Z$, it follows that

$$
\begin{aligned}
g^{(k+1)} - g^* \leq \langle Y^{(k+1)} - \mu \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{A}(X^* - X^{(k+1)}) \rangle \\
+ \langle Y^{(k+1)}, \mathcal{B}(W^* - W^{(k+1)}) \rangle \\
= \langle Y^{(k+1)}, Z - \mathcal{A}X^{(k+1)} - \mathcal{B}W^{(k+1)} \rangle \\
- \mu \langle \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{A}(X^* - X^{(k+1)}) \rangle \\
= \langle Y^{(k+1)}, -E^{(k+1)} \rangle \\
- \mu \langle \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{B}(W^{(k+1)} - W^*) - E^{(k+1)} \rangle,
\end{aligned}
$$

where the last equality holds due to $\mathcal{A}X^{(k+1)} = E^{(k+1)} - \mathcal{B}W^{(k+1)} + Z$. This completes the proof. $\square$

Lemma 3. *Suppose that (X^*, W^*, Y^*) is the optimal solution to the unaugmented Lagrangian function (27). Denote $E^{(k+1)} = \mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z$ and define*

$$\xi_k = \frac{1}{\mu}\|Y^{(k)} - Y^*\|_F^2 + \mu\|\mathcal{B}(W^{(k)} - W^*)\|_F^2.$$

Then, ξ_k decreases in each iteration and satisfies the following relationship:

$$\xi_k - \xi_{k+1} \geq \mu(\|E^{(k+1)}\|_F^2 + \|\mathcal{B}(W^{(k+1)} - W^{(k)}\|_F^2). \tag{33}$$

Proof. Adding the two inequalities in Lemmas 1 and 2 and multiplying through by 2 give that

$$\begin{aligned}
&2\langle Y^{(k+1)} - Y^*, E^{(k+1)}\rangle - 2\mu\langle\mathcal{B}(W^{(k+1)} - W^{(k)}), E^{(k+1)}\rangle \\
&+ 2\mu\langle\mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{B}(W^{(k+1)} - W^*)\rangle \leq 0.
\end{aligned} \tag{34}$$

Using $Y^{(k+1)} = Y^{(k)} + \mu E^{(k+1)}$, the first term of (34) can be rewritten as

$$\begin{aligned}
&2\langle Y^{(k+1)} - Y^*, E^{(k+1)}\rangle \\
&= 2\langle Y^{(k)} + \mu E^{(k+1)} - Y^*, E^{(k+1)}\rangle \\
&= 2\langle Y^{(k)} - Y^*, E^{(k+1)}\rangle + \mu\|E^{(k+1)}\|_F^2 + \mu\|E^{(k+1)}\|_F^2 \\
&= \frac{2}{\mu}\langle Y^{(k)} - Y^*, Y^{(k+1)} - Y^{(k)}\rangle + \frac{1}{\mu}\|Y^{(k+1)} - Y^{(k)}\|_F^2 \\
&\quad + \mu\|E^{(k+1)}\|_F^2.
\end{aligned}$$

Due to the fact that $Y^{(k+1)} - Y^{(k)} = (Y^{(k+1)} - Y^*) - (Y^{(k)} - Y^*)$, we have

$$\begin{aligned}
&2\langle Y^{(k+1)} - Y^*, E^{(k+1)}\rangle \\
&= \frac{1}{\mu}\left(\|Y^{(k+1)} - Y^*\|_F^2 - \|Y^{(k)} - Y^*\|_F^2\right) + \mu\|E^{(k+1)}\|_F^2.
\end{aligned} \tag{35}$$

Now, we rewrite the remaining terms, i.e.,

$$\begin{aligned}
&\mu\|E^{(k+1)}\|_F^2 - 2\mu\langle\mathcal{B}(W^{(k+1)} - W^{(k)}), E^{(k+1)}\rangle \\
&+ 2\mu\langle\mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{B}(W^{(k+1)} - W^*)\rangle,
\end{aligned}$$

where $\mu\|E^{(k+1)}\|_F^2$ is taken from (35). Replacing $W^{(k+1)} - W^* = (W^{(k+1)} - W^{(k)}) + (W^{(k)} - W^*)$ gives

$$\begin{aligned}
&\mu\left(\|E^{(k+1)} - \mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2\right) \\
&+ \mu\|\mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2 \\
&+ 2\mu\langle\mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{B}(W^{(k+1)} - W^*)\rangle.
\end{aligned}$$

With $W^{(k+1)} - W^{(k)} = (W^{(k+1)} - W^*) - (W^{(k)} - W^*)$, this can be converted into

$$\begin{aligned}
&\mu\left(\|E^{(k+1)} - \mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2\right) \\
&+ \mu\left(\|\mathcal{B}(W^{(k+1)} - W^*)\|_F^2 - \|\mathcal{B}(W^{(k)} - W^*)\|_F^2\right).
\end{aligned}$$

Consequently, (34) is equivalent to

$$\frac{1}{\mu}\left(\|Y^{(k+1)} - Y^*\|_F^2 - \|Y^{(k)} - Y^*\|_F^2\right)$$

$$+\mu\left(\|\mathcal{B}(W^{(k+1)} - W^*)\|_F^2 - \|\mathcal{B}(W^{(k)} - W^*)\|_F^2\right)$$

$$+\mu\left(\|E^{(k+1)} - \mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2\right) \le 0,$$

in other words,

$$\xi_k - \xi_{k+1} \ge \mu\left(\|E^{(k+1)} - \mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2\right) \ge 0.$$

Thus, ξ_k decreases.

Further, in order to obtain (33), it only suffices to verify that the term $\langle E^{(k+1)}, \mathcal{B}(W^{(k+1)} - W^{(k)})\rangle \le 0$. In fact, recalling that $W^{(k+1)}$ minimizes $-\mathrm{Tr}(A_l W B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W\rangle$ and $W^{(k)}$ minimizes $-\mathrm{Tr}(A_l W B_l^\top) + \langle Y^{(k)}, \mathcal{B}W\rangle$ in Lemma 2, we can add

$$-\mathrm{Tr}(A_l W^{(k+1)} B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W^{k+1}\rangle$$

$$\le -\mathrm{Tr}(A_l W^{(k)} B_l^\top) + \langle Y^{(k+1)}, \mathcal{B}W^{(k)}\rangle$$

and

$$-\mathrm{Tr}(A_l W^{(k)} B_l^\top) + \langle Y^{(k)}, \mathcal{B}W^{(k)}\rangle$$

$$\le -\mathrm{Tr}(A_l W^{(k+1)} B_l^\top) + \langle Y^{(k)}, \mathcal{B}W^{(k+1)}\rangle$$

to obtain that

$$\langle Y^{(k+1)} - Y^{(k)}, \mathcal{B}(W^{(k+1)} - W^{(k)})\rangle \le 0.$$

Combining $Y^{(k+1)} - Y^{(k)} = \mu E^{(k+1)}$ and $\mu > 0$, we have

$$\mu\langle E^{(k+1)}, \mathcal{B}(W^{(k+1)} - W^{(k)})\rangle \le 0.$$

This completes the proof. □

Based on the above results, the following convergence theorem is established.

Theorem 1. *Under the conditions of Lemmas 1–3, the iterative solution $(X^{(k)}, W^{(k)})$ converges to the optimal solution when k tends to infinity. Namely, $g^k \to g^*$ as $k \to \infty$.*

Proof. From Lemma 3, ξ_k decreases in each iteration and meets

$$\xi_k - \xi_{k+1} \ge \mu(\|E^{(k+1)}\|_F^2 + \|\mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2).$$

Adding all the terms on both sides and rearranging it, we have

$$\sum_{k=1}^{\infty}\left(\mu(\|E^{(k+1)}\|_F^2 + \|\mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2)\right)$$

$$\le \sum_{k=1}^{\infty}(\xi_k - \xi_{k+1}) \le \xi_1 < \infty.$$

This indicates $\|E^{(k+1)}\|_F^2 + \|\mathcal{B}(W^{(k+1)} - W^{(k)})\|_F^2 \to 0$ (as $k \to \infty$). Due to the non-negativity of the Frobenius norm, we know $E^{(k+1)} \to 0$ and $\mathcal{B}(W^{(k+1)} - W^{(k)}) \to 0$

when $k \to \infty$. By the definition of $\mathcal{A}$ and $\mathcal{B}$, it is easy to find $W^{(k+1)} - W^{(k)} \to 0$ and $X^{(k+1)} - X^{(k)} \to 0$. Combining Lemmas 1 and 2, it follows that

$$
\begin{aligned}
g^* - g^{(k+1)} &\leq \langle Y^*, \mathcal{A}X^{(k+1)} + \mathcal{B}W^{(k+1)} - Z \rangle \\
&= \langle Y^*, E^{(k+1)} \rangle \to 0,
\end{aligned}
$$

and

$$
\begin{aligned}
g^{(k+1)} - g^* &\leq \mu \langle \mathcal{B}(W^{(k+1)} - W^{(k)}), \mathcal{B}(W^{(k+1)} - W^*) - E^{(k+1)} \rangle \\
&\quad - \langle Y^{(k+1)}, E^{(k+1)} \rangle \to 0
\end{aligned}
$$

when $k \to \infty$. That is to say, $g^k \to g^*$ as $k \to \infty$. This completes the proof. $\square$

In terms of Theorem 1, it can be seen that the iteration solution to the objective function approximates the optimal value, illuminating the convergence property of the proposed method.

Remark 1. *The GMC-GBTNR model enhances the geometric matrix completion task by incorporating both row-graph and column-graph structures. It advances beyond traditional model-driven methods by integrating structural information with truncated norm details, offering a more comprehensive solution to matrix completion challenges. Unlike the data-driven model, which may lack a solid theoretical underpinning, the GMC-GBTNR stands on a firm theoretical foundation. Its convergence properties, as substantiated in Section 4.3, provide a reliable basis for its effectiveness, making it a robust and theoretically sound approach in the realm of recommendation systems. The GMC-GBTNR framework, while effective, can potentially be elevated by integrating both data-driven and model-driven schemes. Such a hybrid approach may yield even better performance in recommendation tasks due to its comprehensive analysis of data and underlying theoretical models. Nonetheless, crafting a theoretical framework for this hybrid model presents a significant challenge, one that is complex and nuanced, and thus, it remains an area for future research exploration. This endeavor would aim to balance the empirical strengths of data-driven methods with the predictive reliability of model-driven frameworks, carving out new frontiers in the recommendation domain.*

5. Experiments

5.1. Datasets

In our experiments, we employ five specialized datasets focusing on learning resources, as meticulously described in reference [44], detailed as follows:

- Ratings dataset (RA) (https://www.kaggle.com/philippsp/book-recommender-collaborative-filtering-shiny, accessed on 1 September 2023): This learning resource dataset encompasses an extensive compilation of book evaluations, featuring approximately 980,000 ratings for 10,000 titles by 53,424 participants. It showcases the application of collaborative filtering techniques, providing an insightful perspective on algorithm-based book recommendations.
- BX-Book-ratings dataset (BBA) (https://www.kaggle.com/ruchi798/bookcrossing-dataset, accessed on on 1 September 2023): This dataset is a comprehensive aggregation of book assessments, comprising 1,149,780 ratings (both explicit and implicit) contributed by 105,283 users for roughly 340,556 books.
- Ratings-Books dataset (RAB) (https://jmcauley.ucsd.edu/data/amazon/, accessed on on 1 September 2023): It includes a vast array of Amazon product evaluations, last updated in 2018. It comprises a detailed collection of data points such as user, item, rating, and timestamp.
- Related-Article Recommendation dataset (RAR) (https://doi.org/10.7910/DVN/AT4MNE, accessed on on 1 September 2023): It originates from a recommender system utilized in the context of digital libraries and reference management soft-

ware. This dataset is specifically oriented towards literature, featuring 2,663,825 users, 7,224,279 books, and a total of 48,879,167 evaluations.

- LibraryThings dataset (LT) (https://cseweb.ucsd.edu/jmcauley/datasets.html, accessed on on 1 September 2023): It not only provides book ratings but also delves into the social connections among its 70,618 users. It contains 1,387,125 ratings covering 385,251 books. The integration of social interactions and rating correlations in this dataset has yielded optimal results.

These datasets exhibit a high level of sparsity, indicating that many users or learning resources have received only a limited number of ratings. In our experiments, alternative versions of these datasets have been created for comparative analysis. Initially, in each dataset, we selectively focus on users and learning resources that have received a minimum of ten ratings. This approach is strategically implemented to effectively mitigate the significant challenge posed by extensive data sparsity. Then, from the refined version of each dataset, we only choose 3000 users and 3000 items, along with the associated ratings linked to these specific user–item interactions. Table 1 summarizes the statistics of these five real-world datasets with details for the number of users, the number of items, the available graph type, the number of ratings, and the density.

Table 1. Statistics of datasets for evaluation. For the "Graphs", "Users/Items" denotes that both user graph and item graph are used.

Dataset	#Users	#Items	Graphs	#Ratings	Density
RA	3000	3000	Users/Items	16,500	0.0018
BBA	3000	3000	Users/Items	10,000	0.0011
RAB	3000	3000	Users/Items	10,800	0.0012
RAR	3000	3000	Users/Items	8100	0.0009
LT	3000	3000	Users/Items	11,700	0.0013

5.2. Baselines and Experimental Setting

In the experiments, we compare our proposed GMC-GBTNR model to the five existing approaches for matrix completion, detailed as follows:

- IALM [17]: This foundational approach employs the inexact augmented Lagrange multiplier method for precise matrix completion.
- MEN [20]: This strategy incorporates the elastic-net regularization scheme, tailored specifically for matrix recovery.
- TNN [19]: This model is notable for being the first to introduce the truncated nuclear norm method in the realm of matrix completion.
- TNNWRE [21]: Building upon TNN [19], this method enhances it by applying differential weights to the rows of the residual error matrix within an augmented Lagrange function. This modification is designed to expedite the model's convergence.
- GMC [10]: This advanced approach utilizes a deep learning framework, based on a multi-graph convolutional neural network architecture, for effective matrix completion.
- SMC [34]: The method described is a spectral geometric matrix completion approach, specifically designed to account for the underlying geometric or topological relationships among the rows and/or columns.
- FMC [15]: This method extends and enhances SMC [34], adopting a functional perspective tailored for geometric matrix completion. It employs a reduced basis to represent a function on the product space of two graphs, which inherently provides robust regularization.

Each of these methodologies provides a unique perspective on matrix completion, against which we assess the efficacy of our proposed method. On the other hand, in assessing the models, this study employs the root mean square error (RMSE), a widely recognized

metric in the field of recommender systems. This is computed using the formula given as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2},$$

where y_i represents the actual rating value, $\hat{y}_i$ is the predicted rating value, and N denotes the total number of missing values/elements in the matrix.

5.3. Results and Discussion

In our experiments, we conduct a performance comparison across five real-world datasets (RA, BBA, RAB, RAR, and LT) using various matrix completion methods. As shown in Table 2, the evaluation encompasses a range of established approaches such as IALM, MEN, TNN, TNNWRE, and GMC, alongside our novel GMC-GBTNR method. The results clearly demonstrate the efficacy of these methods in handling the complexities of different datasets. For instance, in the RA and BBA datasets, while the GMC method shows impressive results with scores of 0.8823 and 3.2578, respectively, our GMC-GBTNR method outshines the rest by recording the lowest and thus the most favorable scores of 0.8252 and 3.1256, respectively. This suggests a significant enhancement in matrix completion capabilities, highlighting the strengths of our approach in these particular datasets. In particular, in the RA and BBA datasets, the method's superior scores not only demonstrate its effectiveness but also suggest a more nuanced understanding and handling of the inherent complexities within these datasets. This might include better capturing of user–item interactions or more accurately modeling sparse data. In the RAB, RAR, and LT datasets, GMC-GBTNR's consistent outperformance over other methods like TNNWRE indicates its robust adaptability to different data structures and types. These results could be attributed to the method's advanced algorithmic structure, which effectively leverages the truncated Dirichlet norms. The truncated Dirichlet norms in the GMC-GBTNR model play a crucial role in enhancing its matrix completion capabilities. These norms function by prioritizing the significant singular values and corresponding eigenvectors in the data, allowing the model to focus on the most impactful features. This selective emphasis is particularly effective in datasets with complex structures or sparse information, as it filters out noise and less relevant details. By homing in on these key elements, GMC-GBTNR can achieve more accurate and robust predictions, as evidenced by its remarkable performance across varied datasets. This approach illustrates the model's advanced analytical capacity, making it a powerful tool in discerning and leveraging the essential characteristics of data for superior recommendations.

Furthermore, the performance metrics in the remaining datasets—RAB, RAR, and LT—reiterate the superior performance of the GMC-GBTNR method. Across these datasets, it consistently surpasses the other methods, marking a notable achievement in the field of matrix completion. In RAR, for example, where the TNNWRE scores 0.0510, our method advances further with a score of 0.0356. Similarly impressive is its performance in the LT dataset, where it achieves a groundbreaking score of 0.8438, significantly ahead of GMC's 0.8780. These results not only validate the robustness and adaptability of our GMC-GBTNR method but also underscore its potential as a leading solution in matrix completion. By consistently outperforming across a spectrum of datasets, our method establishes a new benchmark in accuracy and efficiency, paving the way for future advancements in this domain.

Table 2. Performance comparison for the five real-world datasets. The performance of our proposed method, GMC-GBTNR, is emphasized with **bold text**, indicating its superior performance across all datasets compared to the other methods.

Method	Dataset				
	RA	BBA	RAB	RAR	LT
IALM [17]	0.9567	3.6793	1.0634	0.0516	0.9278
MEN [20]	0.9313	3.5732	0.9812	0.0498	0.9132
TNN [19]	0.9236	3.4178	1.0416	0.0478	0.8921
TNNWRE [21]	0.9342	3.4322	0.9768	0.0510	0.9126
GMC [10]	0.8823	3.2578	0.9532	0.0452	0.8780
SMC [34]	0.8715	3.2468	0.9456	0.0441	0.8672
FMC [15]	0.8689	3.2319	0.9378	0.0436	0.8543
GMC-GBTNR	**0.8252**	**3.1256**	**0.9216**	**0.0356**	**0.8438**

5.4. Demonstration of Convergence

To illustrate the convergence characteristics of the proposed GMC-GBTNR model, we present the loss curves (of RMSE values) for all five datasets. As depicted in Figure 1, each curve initiates from a specific loss value and exhibits a pronounced decrease, transitioning into a more gradual descent as it nears the 200-epoch mark. This pattern underscores a swift initial improvement during the early covergence phase, followed by a progressive stabilization in convergence rate. Notably, the distinct shapes and levels of the loss curves across the datasets underscore the model's adaptability and consistent convergence behavior despite varying dataset complexities. This demonstration effectively showcases the effectiveness of the GMC-GBTNR model in achieving convergence. It provides compelling empirical evidence that supports and verifies our theoretical findings regarding the convergence properties of the GMC-GBTNR model, as detailed in Section 4.3. The consistency observed between the theoretical predictions and the practical outcomes underlines the reliability and effectiveness of the GMC-GBTNR model in real-world applications.
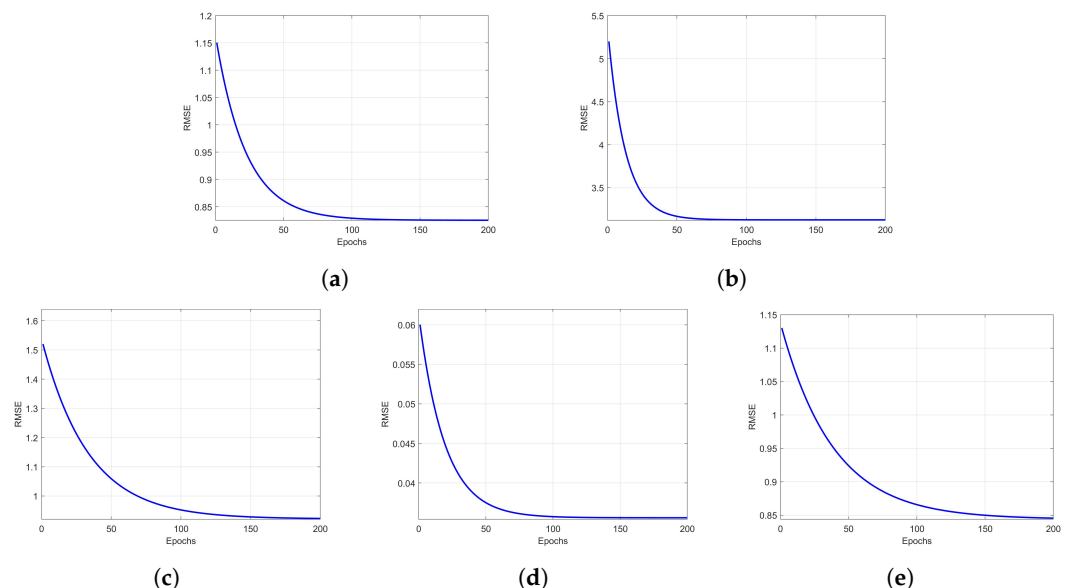
Figure 1. Demonstration for training loss convergence of GMC-GBTNR on five datasets: (**a**) RA, (**b**) BBA, (**c**) RAB, (**d**) RAR, (**e**) LT.

Generally, the convergence property is crucial, as it ensures the stability and reliability of our GMC-GBTNR model. In practical terms, it means that the model will consistently produce accurate and dependable results over time, which is vital for applications in real-world scenarios, as it guarantees that the model's predictions become more accurate and stable over time, regardless of initial conditions or minor changes in the input data.

Theoretically and empirically, this convergence property is key for users who rely on our model for consistent and dependable recommendations, ensuring that the system's performance remains robust and trustworthy in various scenarios.

5.5. Robustness Analysis

In this section, we investigate empirically the robustness of the proposed GMC-GBTNR model. Specifically, we conduct additional experiments introducing varying levels of white Gaussian noise into the rating matrix $X \in \mathbb{R}^{m \times n}$. This noise, represented by $M \sim \mathcal{N}(0, \sigma^2) \in \mathbb{R}^{m \times n}$, was calibrated using $\sigma = p(\max(X) - \min(X))$, where p varies across a set $\{0.01, 0.03, 0.05, 0.08, 0.1\}$. This range of p values allowed us to systematically test the model's performance under different degrees of noise, providing a comprehensive understanding of its resilience to data perturbations.

The experimental results in Table 3 reflect the robustness of the GMC-GBTNR algorithm across five different datasets. Notably, as the noise level increases (with p values ranging from 0.01 to 0.1), the performance of the GMC-GBTNR algorithm exhibits minimal variation, indicating a gradual and stable behavior despite the introduction of noise. This consistent performance across datasets such as RA, BBA, RAB, RAR, and LT suggests that the model maintains its robustness, highlighting its reliability and effectiveness in various conditions. The results are consistent across all datasets, further reinforcing the robust nature of our proposed GMC-GBTNR model.

Table 3. Evaluation of GMC-GBTNR's robustness across different levels of noise.

Case	Dataset				
	RA	BBA	RAB	RAR	LT
GMC-GBTNR (for $p = 0.01$)	0.8210	3.1198	0.9192	0.0326	0.8410
GMC-GBTNR (for $p = 0.03$)	0.8168	3.1172	0.9182	0.0318	0.8392
GMC-GBTNR (for $p = 0.05$)	0.8153	3.1103	0.9145	0.0307	0.8356
GMC-GBTNR (for $p = 0.08$)	0.8121	3.1056	0.9113	0.0301	0.8334
GMC-GBTNR (for $p = 0.1$)	0.8098	3.1012	0.9097	0.0298	0.8312

6. Conclusions and Future Work

In this paper, we introduce a novel geometric matrix completion method tailored for learning resource recommendation, incorporating graph-based truncated norm regularization. This approach not only reflects user-to-user and resource-to-resource relationships but also integrates truncated Dirichlet norms for user and item graphs. This combination enhances the model's capacity to interpret complex data structures and merges the benefits of truncated norm regularization with a thorough analysis of graph relationships. Supported by a comprehensive theoretical analysis and empirical validation across diverse datasets, our GMC-GBTNR model demonstrates reliability and effectiveness in enhancing recommendation systems, thus making a significant contribution to personalized education and enriching our understanding of the dynamics of learner–resource interactions. The GMC-GBTNR model, while effective in certain aspects, faces notable limitations. Its scalability is a primary concern, as the model may struggle with the efficient processing of large datasets, impacting its performance. Additionally, the model's approach of generating row and column graphs without prior knowledge of the rating distribution can be suboptimal, especially in scenarios where such information is critical. Furthermore, the GMC-GBTNR model encounters challenges in handling extremely sparse rating matrices, which can significantly affect its accuracy and reliability in extracting meaningful insights from such data.

For future work, we aim to tackle the scalability in large educational datasets and integrate domain-specific insights into our model. We also plan to explore the potential of large language models (LLMs) to further refine the GMC-GBTNR model, enhancing its relevance and effectiveness in personalized learning environments. We believe that

these future directions promise to extend the impact and applicability of our research in the rapidly evolving field of online education.

Author Contributions: Conceptualization, Y.Y.; Methodology, Y.Y. and S.Y.; Validation, Y.Y., J.S., S.Z. and S.Y.; Formal analysis, Y.Y. and J.S.; Data curation, Y.Y.; Writing—original draft, Y.Y., J.S. and S.Y.; Writing—review & editing, Y.Y., J.S., S.Z. and S.Y.; Visualization, Y.Y. and S.Y. All authors have read and agreed to the published version of the manuscript.

References

1. Thai-Nghe, N.; Drumond, L.; Krohn-Grimberghe, A.; Schmidt-Thieme, L. Recommender system for predicting student performance. *Procedia Comput. Sci.* **2010**, *1*, 2811–2819. [CrossRef]
2. Kaklauskas, A.; Zavadskas, E.K.; Seniut, M.; Stankevic, V.; Raistenskis, J.; Simkevičius, C.; Stankevic, T.; Matuliauskaite, A.; Bartkiene, L.; Zemeckyte, L.; et al. Recommender system to analyze student's academic performance. *Expert Syst. Appl.* **2013**, *40*, 6150–6165.
3. Wu, Z.; Li, M.; Tang, Y.; Liang, Q. Exercise recommendation based on knowledge concept prediction. *Knowl.-Based Syst.* **2020**, *210*, 106481. [CrossRef]
4. Li, M.; Zhang, L.; Cui, L.; Bai, L.; Li, Z.; Wu, X. BLoG: Bootstrapped graph representation learning with local and global regularization for recommendation. *Pattern Recognit.* **2023**, *144*, 109874. [CrossRef]
5. Li, M.; Zhuang, X.; Bai, L.; Ding, W. Multimodal graph learning based on 3D Haar semi-tight framelet for student engagement prediction. *Inf. Fusion* **2024**, *105*, 102224.
6. Ramlatchan, A.; Yang, M.; Liu, Q.; Li, M.; Wang, J.; Li, Y. A survey of matrix completion methods for recommendation systems. *Big Data Min. Anal.* **2018**, *1*, 308–323.
7. Chen, Z.; Wang, S. A review on matrix completion for recommender systems. *Knowl. Inf. Syst.* **2022**, *64*, 1–34. [CrossRef]
8. da Silva, F.L.; Slodkowski, B.K.; da Silva, K.K.A.; Cazella, S.C. A systematic literature review on educational recommender systems for teaching and learning: Research trends, limitations and opportunities. *Educ. Inf. Technol.* **2023**, *28*, 3289–3328. [CrossRef]
9. Dascalu, M.I.; Bodea, C.N.; Mihailescu, M.N.; Tanase, E.A.; Ordoñez de Pablos, P. Educational recommender systems and their application in lifelong learning. *Behav. Inf. Technol.* **2016**, *35*, 290–297.
10. Kalofolias, V.; Bresson, X.; Bronstein, M.; Vandergheynst, P. Matrix completion on graphs. *arXiv* **2014**, arXiv:1408.1717.
11. Monti, F.; Bronstein, M.; Bresson, X. Geometric matrix completion with recurrent multi-graph neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 3697–3707.
12. Berg, R.v.d.; Kipf, T.N.; Welling, M. Graph convolutional matrix completion. *arXiv* **2017**, arXiv:1706.02263.
13. Dong, Y.; Che, H.; Leung, M.F.; Liu, C.; Yan, Z. Centric graph regularized log-norm sparse non-negative matrix factorization for multi-view clustering. *Signal Process.* **2024**, *217*, 109341. [CrossRef]
14. Liu, C.; Wu, S.; Li, R.; Jiang, D.; Wong, H.S. Self-supervised graph completion for incomplete multi-view clustering. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 9394–9406.
15. Sharma, A.; Ovsjanikov, M. Geometric matrix completion: A functional view. *arXiv* **2020**, arXiv:2009.14343.
16. Koren, Y.; Bell, R.; Volinsky, C. Matrix factorization techniques for recommender systems. *Computer* **2009**, *42*, 30–37. [CrossRef]
17. Candès, E.J.; Recht, B. Exact matrix completion via convex optimization. *Found. Comput. Math.* **2009**, *9*, 717. [CrossRef]
18. Srebro, N.; Rennie, J.; Jaakkola, T. Maximum-margin matrix factorization. *Adv. Neural Inf. Process. Syst.* **2004**, *17*.
19. Hu, Y.; Zhang, D.; Ye, J.; Li, X.; He, X. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2117–2130.
20. Li, H.; Chen, N.; Li, L. Error analysis for matrix elastic-net regularization algorithms. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 737–748.
21. Liu, Q.; Lai, Z.; Zhou, Z.; Kuang, F.; Jin, Z. A truncated nuclear norm regularization method based on weighted residual error for matrix completion. *IEEE Trans. Image Process.* **2016**, *25*, 316–330. [CrossRef]
22. Jia, Z.; Jin, Q.; Ng, M.K.; Zhao, X.L. Non-local robust quaternion matrix completion for large-scale color image and video inpainting. *IEEE Trans. Image Process.* **2022**, *31*, 3868–3883. [PubMed]
23. Qin, W.; Wang, H.; Zhang, F.; Wang, J.; Luo, X.; Huang, T. Low-rank high-order tensor completion with applications in visual data. *IEEE Trans. Image Process.* **2022**, *31*, 2433–2448. [PubMed]
24. Masood, K.F.; Tong, J.; Xi, J.; Yuan, J.; Yu, Y. Inductive Matrix Completion and Root-MUSIC-Based Channel Estimation for Intelligent Reflecting Surface (IRS)-Aided Hybrid MIMO Systems. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 7917–7931.

25. Ma, H.; Zhou, D.; Liu, C.; Lyu, M.R.; King, I. Recommender systems with social regularization. In Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011; pp. 287–296.
26. Belkin, M.; Niyogi, P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Adv. Neural Inf. Process. Syst.* **2001**, *14*, 585–591.
27. Belkin, M.; Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **2003**, *15*, 1373–1396. [CrossRef]
28. Rao, N.; Yu, H.F.; Ravikumar, P.K.; Dhillon, I.S. Collaborative filtering with graph information: Consistency and scalable methods. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 2107–2115.
29. Kuang, D.; Shi, Z.; Osher, S.; Bertozzi, A. A harmonic extension approach for collaborative ranking. *arXiv* **2016**, arXiv:1602.05127.
30. Shuman, D.I.; Narang, S.K.; Frossard, P.; Ortega, A.; Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **2013**, *30*, 83–98. [CrossRef]
31. Belkin, M.; Niyogi, P.; Sindhwani, V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *J. Mach. Learn. Res.* **2006**, *7*, 2399–2434.
32. Zhou, D.; Schölkopf, B. A regularization framework for learning from graph data. In Proceedings of the ICML 2004 Workshop on Statistical Relational Learning and Its Connections to Other Fields, Banff, AB, Canada, 8 July 2004; pp. 132–137.
33. Bruna, J.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral networks and locally connected networks on graphs. *arXiv* **2013**, arXiv:1312.6203.
34. Boyarski, A.; Vedula, S.; Bronstein, A. Spectral Geometric Matrix Completion. *Proc. Math. Sci. Mach. Learn.* **2022**, *145*, 172–196.
35. Gao, B.; Absil, P.A. A Riemannian rank-adaptive method for low-rank matrix completion. *Comput. Optim. Appl.* **2022**, *81*, 67–90. [CrossRef]
36. Tsakiris, M.C. Low-rank matrix completion theory via Plücker coordinates. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 10084–10099. [CrossRef] [PubMed]
37. Benzi, K.; Kalofolias, V.; Bresson, X.; Vandergheynst, P. Song recommendation with non-negative matrix factorization and graph total variation. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2439–2443.
38. Cai, J.F.; Candès, E.J.; Shen, Z. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **2010**, *20*, 1956–1982. [CrossRef]
39. Cao, F.; Chen, J.; Ye, H.; Zhao, J.; Zhou, Z. Recovering low-rank and sparse matrix based on the truncated nuclear norm. *Neural Netw.* **2017**, *85*, 10–20. [PubMed]
40. Chung, F.R.; Graham, F.C. *Spectral Graph Theory*; American Mathematical Society: Providence, RI, USA, 1997.
41. Huang, J.; Zhang, F.; Wang, J.; Liu, X.; Jia, J. The perturbation analysis of nonconvex low-rank matrix robust recovery. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**. . [CrossRef]
42. Ye, H.; Li, H.; Cao, F.; Zhang, L. A Hybrid Truncated Norm Regularization Method for Matrix Completion. *IEEE Trans. Image Process.* **2019**, *28*, 5171–5186 [CrossRef]
43. Rockafellar, R.T. *Convex Analysis*; Princeton University Press: Princeton, NJ, USA, 2015.
44. Dien, T.T.; Thanh-Hai, N.; Thai-Nghe, N. An approach for learning resource recommendation using deep matrix factorization. *J. Inf. Telecommun.* **2022**, *6*, 381–398. [CrossRef]