A Geometric Approach to Low-Rank Matrix and Tensor Completion

by

Kenneth Allen

(Under the Direction of Ming-Jun Lai)

Abstract

Data sets arranged as an incomplete matrix or tensor appear frequently in problems such as creating recommendation systems. We introduce a new gradient descent based approach to solving low-rank matrix and tensor completion problems which utilizes maximum volume algorithms by minimizing the norm of Schur compliments. We also prove that the low-rank completions of a partially known matrix may be recovered algebraically by only considering the zero sets of minors containing a known full rank submatrix. We discuss some topological properties of the space of low-rank completions and introduce a relationship between the zeroth Betti number of the space of low-rank matrix or tensor completions and the first Betti number of the union of the space of completions and the space of low-rank matrices or tensors. Additionally, we give conditions for partially known tensors with certain structures to have unique low-rank, low-border-rank, or low-multilinear-rank completion.

A greedy version of the maximum volume algorithm for finding dominant submatrices is presented which may permute more than one row at each step and can improve on the processing time of the original maximum volume algorithm. Moreover, we introduce an upper bound on the number of $r \times r$ dominant submatrices of almost all $m \times n$ matrices in terms of the independence number of the Cartesian product of Johnson graphs $\alpha(J_{m,r} \Box J_{n,r})$. We conjecture that this bound is the essential supremum of the function which counts the number of $r \times r$ dominant submatrices.

We show that a maximum volume skeleton decomposition method may be used as a scalable alternative to the singular value decomposition as a low-rank approximation step in the dynamic mode decomposition. We also show that the maximum volume skeleton decomposition may be utilized to compress large sets of tokamak edge plasma simulation data.

Index words: [Data compression, Johnson graph, Matrix completion, Maximum volume, Plasma simulation, Tensor completion]

A Geometric Approach to Low-Rank Matrix and Tensor Completion

by

Kenneth Allen

B.S., University of Massachusetts Lowell, 2017

A Dissertation Submitted to the Graduate Faculty of the
University of Georgia in Partial Fulfillment of the Requirements for the Degree.

Doctor of Philosophy

Athens, Georgia

2021

A Geometric Approach to Low-Rank Matrix and Tensor Completion

by

Kenneth Allen

Major Professor: Ming-Jun Lai

Committee: Weiwei Hu
Lin Mu
Alexander Petukhov

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2021

Dedication

This dissertation is dedicated to my late father Bradford Allen who took every opportunity to express how much he loved me. I love you, dad.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE-SC0014664.

I'm deeply indebted to my advisor Ming-Jun Lai who has consistently provided an incredible amount of support, and without whom my success would not have been possible. I would also like to extend my deepest gratitude to Lin Mu who helped me immensely by supporting me with a research assistantship and introducing me to those at ORNL. I wish to thank Sebastian De Pascuale and David Green at ORNL for their excellent collaboration.

I'm extremely grateful to my mother Elaine Allen for all the support and encouragement throughout the years. I also wish to thank James Calhoun for the invaluable emotional support during this process. I am grateful to Nolan Schock and Benjamin Tighe for the helpful mathematical conversations. I would also like to acknowledge Grant Moyer for assistance with Tikz pictures in this dissertation. Many thanks to all my other friends and family for their support as well.

Contents

List of Tables

List of Figures

List of Algorithms

Nomenclature

Elements

M_Ω A partially known matrix with known entries M_{ij} in positions $(i, j) \in \Omega$ and zeros elsewhere

$M_{I,J}$ The submatrix of M with entries in positions $I \times J$ for sets of indices $I \subset [m]$, $J \subset [n]$

M_{ij} The entry of the matrix M in position (i, j)

$\square$ A missing element of a matrix or tensor

T_Ω A partially known tensor with known entries T_{ijk} in positions $(i, j, k) \in \Omega$ and zeroes elsewhere

T_{ijk} The entry of the third order tensor T in position (i, j, k)

$\hat{X}$ A dominant submatrix of a matrix X

$X_{\max}$ A maximum volume submatrix of a matrix X

Functions

$\alpha(G)$ The independence number of the graph G

$B(\Omega)$ The binary matrix or tensor with a one in positions in Ω zeros elsewhere

$\deg(V)$ The degree of the algebraic variety V

$\det(M)$ The determinant of the matrix M

$\mathrm{diag}(x_1, \ldots, x_n)$ The diagonal matrix with diagonal entries equal to $x_1, \ldots, x_n$

$\mathrm{grank}(T)$ The generic rank of the tensor T

$h_i(V)$ The ith Betti number of the topological space V

$\mu\mathrm{rank}(T)$ The multilinear rank of the tensor T

$\|M\|_*$ The nuclear norm of the matrix M, equal to the sum of the singular values of M

$\|M\|_2$ The spectral norm of the matrix M, equal to the largest singular value of M

$\|M\|_{\max}$ The max norm of the matrix M

M^* The conjugate transpose of the matrix M

$M^\dagger$ The Moore–Penrose pseudoinverse of the matrix M

$M^\top$ The transpose of the matrix M

$M^{-\top}$ The inverse of the transpose of the matrix M

$\Phi_\Omega(M)$ The restriction of the projection map P_Ω to $\overline{\mathcal{M}}_r$

$P_\Omega(X)$ The projection of the matrix or tensor X that sets entries in positions not in Ω to zero

$P_M(X)$ The orthogonal projection operator onto the column space of the matrix M given by $MM^\dagger X$

$P_{(i,j)}(M)$ The map which gives M_{ij}, the entry of the matrix M in position (i, j)

$P_{\mathcal{M}_r}(M)$ A closest rank r approximation of the matrix M

$\overline{\mathrm{rank}}(T)$ The border rank of the tensor T

$\mathrm{rank}(X)$ The rank of the matrix or tensor X

$\sigma_i(M)$ The ith singular value of a matrix M

S_A The Schur complement of the matrix M with respect to the submatrix A

$\mathrm{tr}(M)$ The trace of the matrix M

$\mathrm{vec}(M)$ The vectorization of the matrix M

$\mathrm{vol}(M)$ The volume of the matrix M, equal to the absolute value of the determinant of M

$|X|$ The cardinality of the set X, or is the absolute value of the real or complex number X

$\chi(V)$ The Euler characteristic of the topological space V

$\langle X, Y\rangle$ The inner product $\mathrm{tr}(X^*Y)$ of the matrices X and Y

$\|X\|$ The Euclidean norm of the vector, matrix, or tensor X

Sets

$\mathcal{A}_\Omega$ The linear affine space of all completions of M_Ω or T_Ω

$\mathbb{A}^d$ The affine space of real numbers $\mathbb{R}^d$ or complex numbers $\mathbb{C}^d$

$B_\epsilon(X)$ The ball centered at X with radius ϵ

$\mathbb{C}^{m\times n}$ The set of m by n complex matrices

$E(G)$ The set of edges in the graph G

$G\square H$ The graph Cartesian product of the graphs G and H

$H_i(V)$ The ith homology group of the topological space V

$J_{m,r}$ The Johnson graph whose vertices are the subsets size k of $[m]$, with an edge connecting two vertices if their intersection contains $r-1$ elements

$[m]\times[n]$ The set of positions (i,j) such that $1\le i\le m$ and $1\le j\le n$

$\mathcal{M}_r$ The set of $m \times n$ matrices with rank equal to r

$\overline{\mathcal{M}_r}$ The set of $m \times n$ matrices with rank at most r

$M_{m\times n}$ The set of matrices with m rows and n columns over $\mathbb{R}$ or $\mathbb{C}$

$[n]$ The set of integers $\{1, \ldots, n\}$

Ω The set of observed positions of a partially known matrix or tensor

$\mathbb{R}^{m\times n}$ The set of m by n real matrices

$\hat{\sigma}_r$ The set of tensors in $U \otimes V \otimes W$ with border rank at most r

σ_r The set of tensors in $U \otimes V \otimes W$ with rank at most r

$\hat{S}ub_r$ The set of tensors in $U \otimes V \otimes W$ such that each component of the multilinear rank is at most r

$V(G)$ The set of vertices in the graph G

V^* The dual of the vector space V

Chapter 1

Low-Rank Matrix Completion

1.1 Introduction

In 2006, Netflix announced a competition with a grand prize of one million dollars. The problem was, given data on user's movie ratings, create a recommendation system which would suggest films users would be likely to watch and enjoy. The grand prize would be given out to anyone who could improve Netflix's existing algorithm by ten percent. This is an example of a data completion problem, where we have an incomplete set of data, and we would like to complete that set of data using the known data as best as possible.

It is often useful to encode partially known data in a matrix. For Netflix's problem, on one axis we index the users, and on the other axis we index the movies. In the corresponding entry between a user and a movie, we enter the user's rating of the movie. The result is a partially known matrix, and the goal is to fill in the matrix as best as possible. In other words, Netflix would like to predict how users will rate movies that they have not watched so they can suggest movies they think the user will enjoy. This problem sparked an interest among mathematicians in studying the matrix completion problem.

A first approach to the matrix completion problem is to define a number of features of the data which we can use to calculate intermediate connections between users and movies. For movies, we can use genres such as *fantasy* and *mystery* as features. We assign strengths between users and features, and strengths between features and movies to decide whether or not we should recommend a movie.

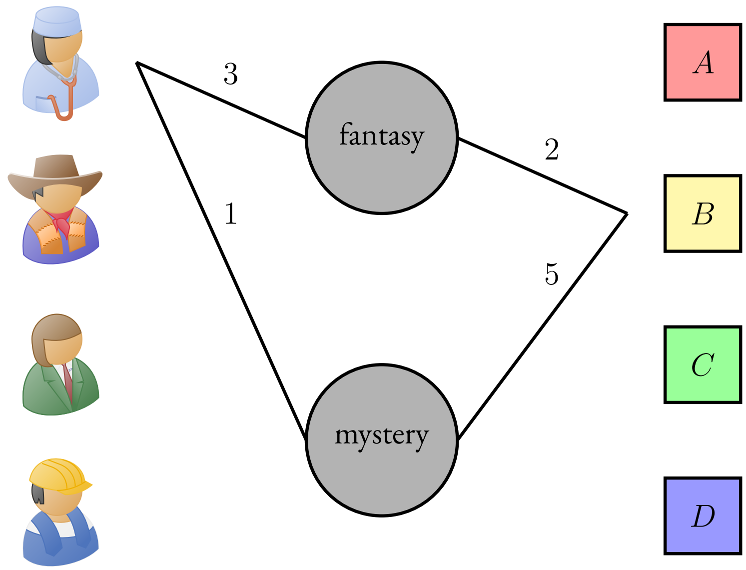


Figure 1.1: Example of strengths between users, features, and movies

To determine if a user would like a movie, we multiply the strengths between the user-feature score and feature-movie score, and sum over all features. In the example shown in Figure 1.1, the first user would be given a score of $3 \cdot 2 + 1 \cdot 5 = 11$ as a prediction for how much they would enjoy movie B. There are multiple ways we can get this data. For example, Netflix had at one point sent users surveys to determine which movie genres they enjoyed the most.

We can represent user-feature and feature-movie data as two matrices as shown in Figure 1.2. To obtain the corresponding user-movie rating, we simply multiply these two matrices together. The issues with this method are that it may be unrealistic to gather this data from all users or all movies. Moreover, these human-created features may not be the most general way to represent the data.

We may instead reverse the problem. We measure user enjoyment of some movies through data such as ratings as in Figure 1.3. Then, given some incomplete data on user-movie enjoyment, we produce two factor matrices U and V such that the corresponding entries in the multiplication UV agrees with the known data. If we can find such matrices then we may predict any user-movie enjoyment score by observing the corresponding entry in UV.

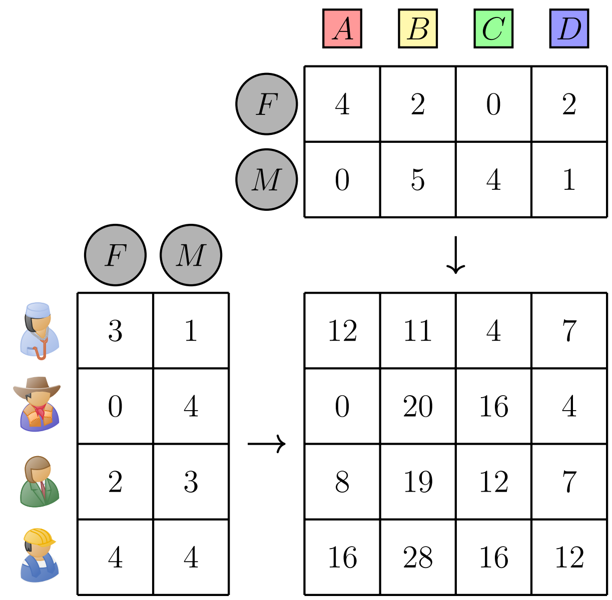


Figure 1.2: User-Movie rating prediction matrix obtained through matrix multiplication.

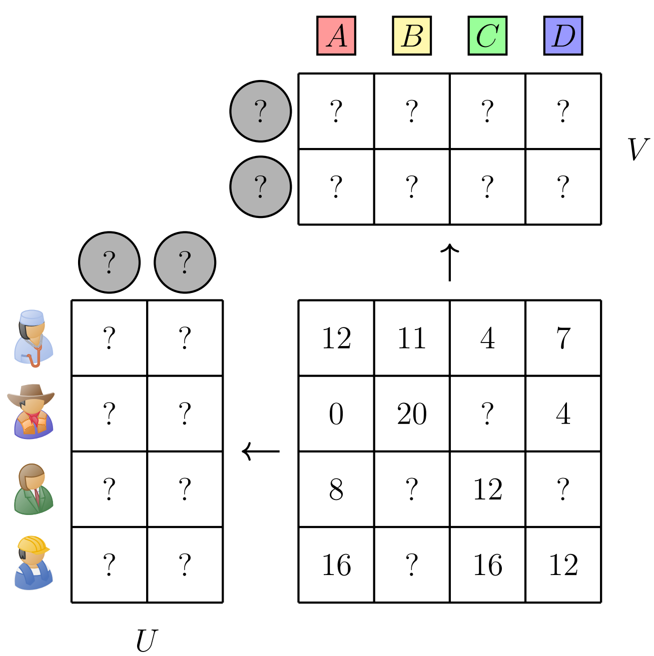


Figure 1.3: User-Movie prediction matrix may be obtained by finding factor matrices

Note that the product UV will be low-rank, depending on the size of U and V. In particular, if we assume there are r features which describe the data, U is $m \times r$, and V is $r \times n$ the resulting matrix will be $m \times n$ and have rank at most r. Therefore, we don't necessarily need to find explicit factor matrices U and V, we just need to find a rank r matrix M with entries equal to the given known entries.

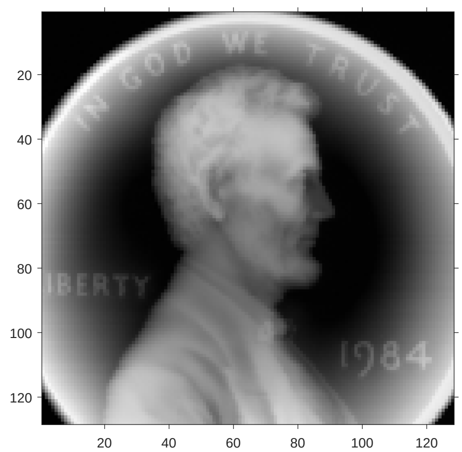
Why should we assume that a given data set is part of a low rank matrix? Not all matrices will be close to a low-rank matrix. For example, a random Gaussian matrix will not be close to a low rank matrix with very high probability (Edelman, 1988). That being said, close to low-rank matrices appear very frequently throughout the sciences, which means they are far from average. In fact, it has been shown that under certain assumptions, if a matrix of data is drawn in a consistent way, it must be close to low-rank (Udell & Townsend, 2019). This suggest that the low-rank structure of large data sets is not merely coincidental, but rather is a universal feature.

Figure 1.4: Mondrian, Piet. *Composition with Red, Blue, and Yellow* 1930, may be expressed as the sum of four rank one matrices. Black represents zeros.

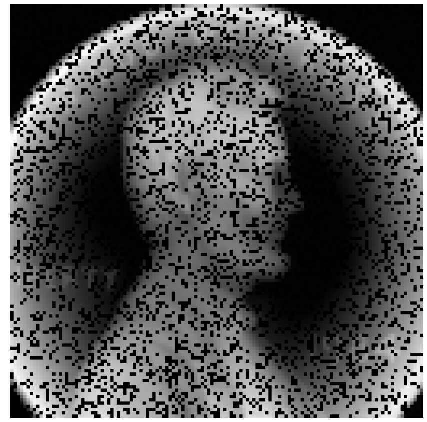
Images can be useful for testing low-rank matrix completion techniques. A rectangular grayscale image with a resolution of $m \times n$ pixels may be represented as an $m \times n$ matrix. Each position in the matrix corresponds to a pixel, with the entry in that position corresponding to the intensity.

We may apply low-rank matrix completion techniques to recover the missing elements of partially complete images such as Figure 1.5b under the assumption that they are close to a low-rank matrix. In Section 2.7, we introduce a novel low-rank matrix completion technique.

Factoring a matrix into a product of two smaller matrices also can also be used to compress data. If M is an $m \times n$ array of data that has rank r, then we may find matrices U and V such that $M = UV$ where U is an $m \times r$ matrix and V is an $r \times n$ matrix. mn entries need to be saved to store M, but only $(m+n)r$ entries need to be saved to store U and V. If r is much smaller than m and n, then it is more efficient to store U and V than it is to store M. In general, M is not low-rank, but it is common for data sets to be



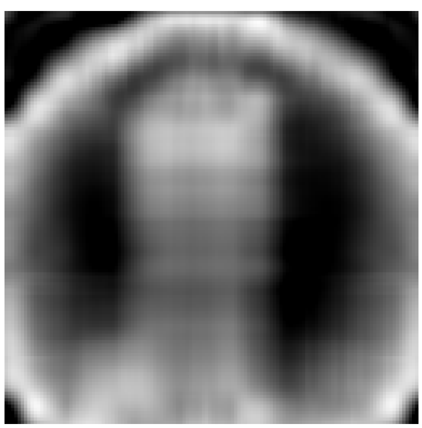
(a) 128×128 penny picture, pixels correspond to integers from 0 to 255



(b) 25% of image deleted uniformly at random

Figure 1.5: Is it possible to recover Figure 1.5a from fig. 1.5b?

close to low rank matrices. That is, we may approximate $M \approx UV$ with small error. For example, see Figure 1.6. In Section 3.3, we apply data compression techniques to plasma simulation data.



(a) Rank 5 approximation

(b) rank 35 approximation

Figure 1.6: Low rank approximations of 128×128 Figure 1.5a

1.2 Mathematical Preliminaries

We now express the matrix completion problem in more formal mathematical terms. Let $M_{m\times n}$ denote the space of matrices with m rows and n columns over the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. Let $[n]$ denote the set of integers $\{1,\ldots,n\}$, and let $[m]\times[n]$ denote the set of positions (i,j) such that $1\leq i\leq m$ and $1\leq j\leq n$.

Definition 1. *Given a set of observed elements $\{M_{ij}\}$, where M_{ij} is in position (i,j) of a partially known $m\times n$ matrix, we denote by $\Omega\subset[m]\times[n]$ the set containing the positions of the known entries M_{ij}, and we refer to Ω as a* mask. *In other words, if M_{ij} is a known element, then $(i,j)\in\Omega$. We denote by $|\Omega|$ the cardinality of the set Ω, then $|\Omega|$ is equal to the total number of known elements M_{ij}. Let $B(\Omega)$ denote the binary matrix where the entry in position (i,j) is equal to one if $(i,j)\in\Omega$, and is equal to zero otherwise. Because Ω and $B(\Omega)$ may be used interchangeably, we also refer to $B(\Omega)$ as a mask.*

It is desirable to use as few features possible to represent the data, so we wish to minimize the rank of the completed matrix. We may now express the matrix completion problem as finding a solution to the non-convex minimization problem

$$\min_{X\in M_{m\times n}} \operatorname{rank}(X)$$

$$\textit{s.t. } X_{ij}=M_{ij}\ \forall(i,j)\in\Omega$$

Let $M_\Omega\in M_{m\times n}$ denote the partially known $m\times n$ matrix with known entries M_{ij} in entry $(i,j)\in\Omega$, and zeros in unknown entries. We often denote unknown elements of M_Ω with an empty square $\square$ instead of zero. There may be finitely many, infinitely many, or zero ways to complete M_Ω into a rank r matrix. Consider the following incomplete matrices.

Example 2. *Let*

$$M_\Omega = \begin{bmatrix} 3 & 2 & 1 & \square \\ 4 & 3 & \square & 1 \\ 5 & \square & 3 & 1 \\ \square & 5 & 4 & 1 \end{bmatrix}.$$

Then M_Ω has the unique rank two completion

$$M = \begin{bmatrix} 3 & 2 & 1 & 1 \\ 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 1 \\ 6 & 5 & 4 & 1 \end{bmatrix}.$$

If we change the known entries in the first column of M_Ω in Example 2, we have the following example.

Example 3. *Let*

$$M_\Omega = \begin{bmatrix} 4 & 2 & 1 & \square \\ 5 & 3 & \square & 1 \\ 6 & \square & 3 & 1 \\ \square & 5 & 4 & 1 \end{bmatrix}.$$

Then M_Ω has exactly two rank two completions, which are

$$M_1 = \begin{bmatrix} 4 & 2 & 1 & 1 \\ 5 & 3 & 2 & 1 \\ 6 & 4 & 3 & 1 \\ 7 & 5 & 4 & 1 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 4 & 2 & 1 & -2/3 \\ 5 & 3 & 21/8 & 1 \\ 6 & 39/11 & 3 & 1 \\ 26/3 & 5 & 4 & 1 \end{bmatrix}.$$

To verify that M_1 and M_2 are the only rank two completions of M_Ω, note that a matrix has rank at most r if and only if all $(r+1) \times (r+1)$ minors vanish. Consider the system of equations in four variables obtained by setting all 3×3 minors of the matrix

$$M(x,y,z,w) = \begin{bmatrix} 4 & 2 & 1 & x \\ 5 & 3 & y & 1 \\ 6 & z & 3 & 1 \\ w & 5 & 4 & 1 \end{bmatrix}$$

equal to zero. This gives a system of 16 degree two and degree three polynomials. One can verify using a computer algebra system that M_1 and M_2 are the only two solutions to this system of equations.

We now introduce the space of fixed rank r matrices. Let

$$\mathcal{M}_r = \{X \in M_{m\times n} \mid \text{rank}(X) = r\}$$

denote the space of $m\times n$ rank r matrices over $\mathbb{R}$ or $\mathbb{C}$, where $r \leq \min(m,n)$. Then $\mathcal{M}_r$ is a $(m+n)r-r^2$ dimensional manifold (Guillemin & Pollack, 2010). Note that $\mathcal{M}_r$ is not a closed set. In particular, we may approximate any low-rank matrix as the limit of a sequence of high-rank matrices, but we may not approximate high-rank matrices as the limit of a sequence of low-rank matrices. So the closure of $\mathcal{M}_r$ is the space of matrices with rank at most r, which we denote

$$\overline{\mathcal{M}}_r = \{X \in M_{m\times n} \mid \text{rank}(X) \leq r\}.$$

Because the closure operation does not change the dimension of a manifold the dimension of $\overline{\mathcal{M}}_r$ is also equal to $(m+n)r - r^2$, .

We now recall some notions from algebraic geometry.

Definition 4. *Let* $\mathbb{A}^d$ *denote the* d *dimensional* affine space $\mathbb{R}^d$ *or* $\mathbb{C}^d$*. A closed set* $V \subset \mathbb{A}^d$ *is called an* algebraic variety *if it is the zero set of a set of a system of polynomial equations. An algebraic variety* V *is* generated *by the polynomials* $f_1, f_2, \ldots, f_N$ *if*

$$V = \{x \in \mathbb{A}^d \mid f_1(x) = f_2(x) = \cdots = f_N(x) = 0\}.$$

The Zariski closure *of a set* $S \subset \mathbb{A}^d$ *is the smallest algebraic variety* V *which contains* S*.*

The Zariski closure of S may be expressed as the intersection of all algebraic varieties which contain S. Because a matrix M has rank at most r if and only if all $(r+1) \times (r+1)$ minors of M vanish, $\overline{\mathcal{M}_r}$ is an algebraic variety generated by the set of all $(r+1) \times (r+1)$ minors of an $m \times n$ matrix. Therefore $\overline{\mathcal{M}_r}$ is the Zariski closure of $\mathcal{M}_r$ because it is the closure of $\mathcal{M}_r$ in the analytic sense and it is an algebraic variety. $\overline{\mathcal{M}_r}$ is also sometimes referred to as the determinantal variety.

Definition 5. *An algebraic variety* V *is called* irreducible *if it cannot be expressed as the union of two proper sub-varieties. That is, an algebraic variety* V *is irreducible if it cannot be written as* $V = V_1 \cup V_2$ *for proper subvarieties* $V_1 \subset V$ *and* $V_2 \subset V$*.*

It is known that $\overline{\mathcal{M}_r}$ is an irreducible variety (Lai & Varghese, 2017). It can also be shown that the set of singular points of $\overline{\mathcal{M}_r}$ is the set of matrices with rank at most $r-1$, that is, it is the set $\overline{\mathcal{M}_{r-1}} \subset \overline{\mathcal{M}_r}$. To verify this note that the partial derivatives of the all $(r+1) \times (r+1)$ minors are equal to zero exactly on the set $\overline{\mathcal{M}_{r-1}}$.

Define the orthogonal projection operator $P_\Omega : M_{m\times n} \to M_{m\times n}$ such that $P_\Omega(X)$ fixes entry X_{ij} if $(i,j) \in \Omega$, and sets X_{ij} equal to zero if $(i,j) \notin \Omega$.

Example 6. *Let* $\Omega = \{(1,1),(2,2)\}$*. Then*

$$P_\Omega\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix}.$$

Given a partially known matrix M_Ω, a matrix X is called a *completion* of M_Ω if $P_\Omega(X) = M_\Omega$. Given M_Ω, let $\mathcal{A}_\Omega = P_\Omega^{-1}(M_\Omega)$ be the linear variety of all completions of M_Ω. In other words,

$$\mathcal{A}_\Omega = \{X \in M_{m\times n} \mid P_\Omega(X) = M_\Omega\}.$$

Because $\mathcal{A}_\Omega$ is a linear variety, it is irreducible. Also because $\overline{\mathcal{M}}_r$ is the space of matrices with rank at most r, and $\mathcal{A}_\Omega$ is the space of completions of M_Ω, Then finding a rank at most r completion M of M_Ω is equivalent to finding a point $M \in \mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$.

The *singular value decomposition* is a matrix decomposition which is useful for calculating low-rank approximations.

Definition 7. *Given an $m \times n$ matrix X, the* singular value decomposition *of X is*

$$X = U\Sigma V^*$$

where U is the $m \times m$ matrix of left-singular vectors *which are the eigenvectors XX^*, V is the $n \times n$ matrix of* right-singular vector *which are the eigenvectors of X^*X, and Σ is the $m \times n$ diagonal matrix of* singular values, *which are the square root of the eigenvalues of XX^* or X^*X.*

Suppose $n \leq m$. Then the singular values of X are denoted $\sigma_i(X)$ for $1 \leq i \leq n$, and are ordered such that $\sigma_i(X) \leq \sigma_{i+1}(X)$ for all i. Let $\|X\|_2 = \sigma_1(X)$ denote the *spectral norm* of the matrix X. The singular values of X are directly related to the rank of X, they give the error of X to a closest low-rank approximation.

Theorem 8. *(Eckart & Young, 1936) $\sigma_{r+1}(X)$ is the error of X to a closest rank r approximation of X in the spectral norm. That is, if X_r is a closest rank r approximation of X in the spectral norm, then*

$$\|X - X_r\|_2 = \sigma_{r+1}(X).$$

In particular, if $\operatorname{rank}(X) = r$, then $\sigma_i(X) = 0$ for all $r + 1 \leq i \leq n$.

We may use the singular value decomposition to explicitly calculate a closest rank r approximation of a matrix X.

Theorem 9. *(Eckart & Young, 1936) Given an $m \times n$ matrix X, let*

$$X = U\Sigma V^*$$

be the singular value decomposition of X. Let U_r be the first r columns of U, let V_r be the first r columns of V, and let Σ_r be the $r \times r$ diagonal matrix of the first r singular values of X. Then

$$X_r = U_r \Sigma_r V_r^*$$

is a closest rank r approximation of X in the spectral norm.

Let $\operatorname{tr}(X) = \sum_i X_{ii}$ denote the trace of X. Let $\|X\|$ denote the Euclidean norm of an $m \times n$ matrix X, where

$$\|X\|^2 = \operatorname{tr}(X^*X) = \sum_{i=1}^{m} \sum_{j=1}^{n} |X_{ij}|^2 .$$

In terms of the singular values of X,

$$\|X\|^2 = \sum_{i=1}^{n} \sigma_i(X)^2.$$

It is often desirable to numerically approximate a point $M \in \mathcal{A}_\Omega \cap \overline{\mathcal{M}_r}$ by a sequence $\{X_k\}$ such that $\|X_k - M\|$ goes to zero as k goes to ∞.

Given a grayscale image on a computer, each pixel can be assigned a number corresponding to the intensity of the shade of gray, often from 0 (black) to 255 (white). We use the *peak signal to noise ratio* to measure the quality of image approximations.

Definition 10. *Given an $m \times n$ matrix X with entries ranging from 0 to 255, and an $m \times n$ approximation Y of X, let $E = \frac{1}{mn} \|X - Y\|^2$ be the mean square error of Y with respect to X. The* peak signal to noise ratio*, or* PSNR*, is defined as*

$$\mathrm{PSNR} = 20 \cdot \log_{10} \left(\frac{255}{\sqrt{E}} \right).$$

The larger the peak signal to noise ratio, the higher the quality Y is as an approximation of X. For lossy image compression methods, a PSNR of at least 30 may be considered good.

1.3 Alternating Projection

The alternating projection method has proven to be an effective way to find intersection points between two manifolds (Lewis & Malick, 2008). As the name suggests, starting with an initial guess, we alternate between projecting onto each manifold obtaining successive approximations of a point in the intersection. While most extensively studied in application to find intersections of convex sets, alternating projection methods have also been applied to find intersections non-convex sets. Some issues that may occur are that the projection itself may be difficult to compute. Moreover, a projection onto a non-convex set may not be single valued, and if the initial guess is not good enough, the alternating projection algorithm may not converge.

Given a partially known matrix M_Ω, recall that $\mathcal{A}_\Omega \cap \mathcal{M}_r$ is the set of rank r completions of M_Ω. The alternating projection method may be used to find a rank at most r completion of M_Ω by alternating between projections onto $\mathcal{A}_\Omega$ and $\mathcal{M}_r$.

Note that $\mathcal{A}_\Omega$ is convex, as it is a linear variety, so for all $X \in M_{m\times n}$ there exists a unique projection of X onto $\mathcal{A}_\Omega$. Let $P_{\mathcal{A}_\Omega} : M_{m\times n} \to \mathcal{A}_\Omega$ denote the projection map onto $\mathcal{A}_\Omega$. Then $P_{\mathcal{A}_\Omega}(X)$ is simply calculated by setting all of the entries X_{ij} with positions in Ω to the corresponding known entries M_{ij} in M_Ω. In other words, let $Y = P_{\mathcal{A}_\Omega}(X)$, then

$$Y_{ij} = \begin{cases} M_{ij} & \text{if } (i,j) \in \Omega \\ X_{ij} & \text{if } (i,j) \notin \Omega \end{cases}$$

On the other hand, $\mathcal{M}_r$ is not convex for $r > 0$, so a given $X \in M_{m\times n}$ may not have a unique projection onto $\mathcal{M}_r$. However, if $\operatorname{rank}(X) > r$, then a closest rank r projection may be easily calculated using the singular value decomposition (SVD).

Let $X = U\Sigma V^*$ be the SVD of an $m \times n$ matrix X, where U is an $m \times m$ orthogonal matrix of left-singular vectors, V is an $n \times n$ orthogonal matrix of right-singular vectors, and Σ is an $m \times n$ diagonal matrix of singular values.

Let Σ_r be the $r \times r$ diagonal matrix with diagonal entries equal to the r largest singular values of X, let U_r be the $m \times r$ matrix with columns consisting of the first r left-singular vectors of X, and let V_r be the $n \times r$ matrix with columns consisting of the first r right-singular vectors of X. Then if $\sigma_i(X) > 0$ for all $1 \leq i \leq r$, define the map

$$P_{\mathcal{M}_r}(X) = U_r \Sigma_r V_r^*.$$

Theorem 11. *(Eckart & Young, 1936) $P_{\mathcal{M}_r}(X)$ is a closest rank r approximation of X with respect to any unitary invariant norm. In other words,*

$$P_{\mathcal{M}_r}(X) = \underset{X_r \in \mathcal{M}_r}{\arg\min} \; \|X - X_r\| \,.$$

Moreover, $P_{\mathcal{M}_r}(X)$ is the unique closest rank r approximation of X if $\sigma_r(X) \neq \sigma_{r+1}(X)$.

Let us suppose $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is nonempty. Then given and initial guess X_0 and a tolerance ϵ, the alternating projection Algorithm 1 can be stated as follows.

Algorithm 1: Alternating Projection (Lai & Varghese, 2017)

Input: partially known matrix M_Ω, initial guess X_0, stopping criterion
Result: X_k an approximation of a rank r completion of M_Ω
for $k = 1, \ldots$ **do**
  $Y_k = P_{\mathcal{M}_r}(X_{k-1})$;
  $X_k = P_{\mathcal{A}_\Omega}(Y_k)$;

We require some stopping criterion as an input. For example, we could fix a tolerance ϵ, and loop until $\|X_k - X_{k-1}\| < \epsilon$. Alternatively, if we only want to run the algorithm for a certain number of iterations N, we could loop for $k = 1, \ldots, N$.

Let $T_{\mathcal{A}_\Omega}(M)$ denote the tangent space of $\mathcal{A}_\Omega$ at point M, and let $T_{\mathcal{M}_r}(M)$ denote the tangent space of $\mathcal{M}_r$ at point M. If $M \in \mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$, and $T_{\mathcal{A}_\Omega}(M) \cap T_{\mathcal{M}_r}(M) = \{0\}$, then Algorithm 1 converges to M linearly (Lai & Varghese, 2017).

Example 12. *We use the alternating projection method to recover missing elements of a picture. Consider the picture of the United States penny, Figure 1.5. The image has a resolution of* 128×128 *pixels, and each pixel corresponds to an integer from* 0 *to* 255*.*

Because each entry is bounded from 0 *to* 255*, then given a partially known image* M_Ω*, the set of completions also has bounded entries. That is,*

$$\mathcal{A}_\Omega = \{X \in M_{m\times n} \mid P_\Omega(X) = M_\Omega, 0 \leq X_{ij} \leq 255\}.$$

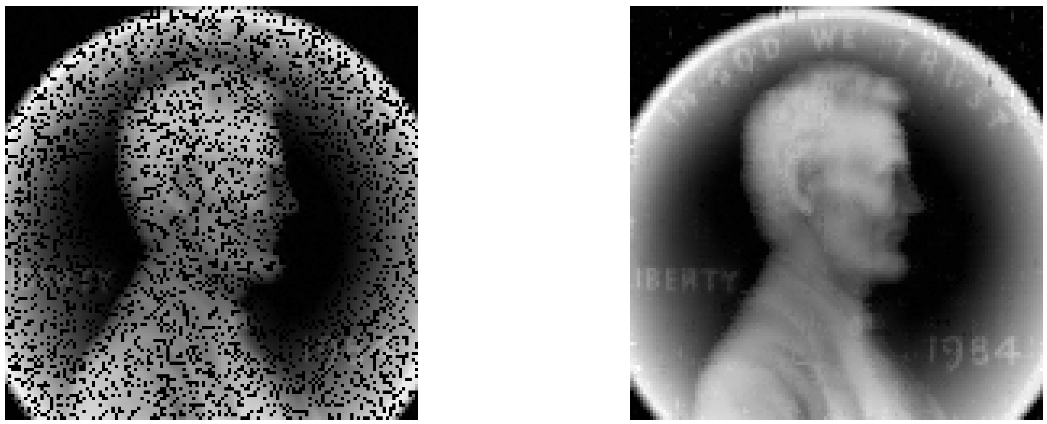
This changes the projection map $P_{\mathcal{A}_\Omega}$*. In particular, if* $Y = P_{\mathcal{A}_\Omega}(X)$*, then*

$$Y_{ij} = \begin{cases} M_{ij} & \textit{if } (i,j) \in \Omega \\ X_{ij} & \textit{if } (i,j) \notin \Omega \textit{ and } 0 \leq X_{ij} \leq 255 \\ 0 & \textit{if } (i,j) \notin \Omega \textit{ and } X_{ij} < 0 \\ 255 & \textit{if } (i,j) \notin \Omega \textit{ and } X_{ij} > 255 \end{cases}$$

In other words, at each projection step onto $\mathcal{A}_\Omega$*, we set all entries in known positions* (i,j) *equal to the known entry* M_{ij}*, and we set all entries greater than* 255 *or less than* 0 *equal to* 255 *and* 0 *respectively.*

We delete 25% *of Figure 1.5 uniformly at random, and attempt to recover the image using Algorithm 1 under the assumption that the rank of the true image is* 18*.*

We can see in Figure 1.8 that the singular values of Figure 1.5 taper off at around $r = 18$*, and the singular values of the recovered picture closely matches the singular values of the original picture.*



(a) 25% of Figure 1.5 deleted uniformly at random

(b) Recovered from Figure 1.7a with algorithm 1

Figure 1.7: Figure 1.5 recovered with 250 iterations of the alternating projection method assuming the rank of the original image is 18. Initial guess is known entries of the image with unknown entries set to zero. Recovery has a peak signal to noise ratio of 34.91.

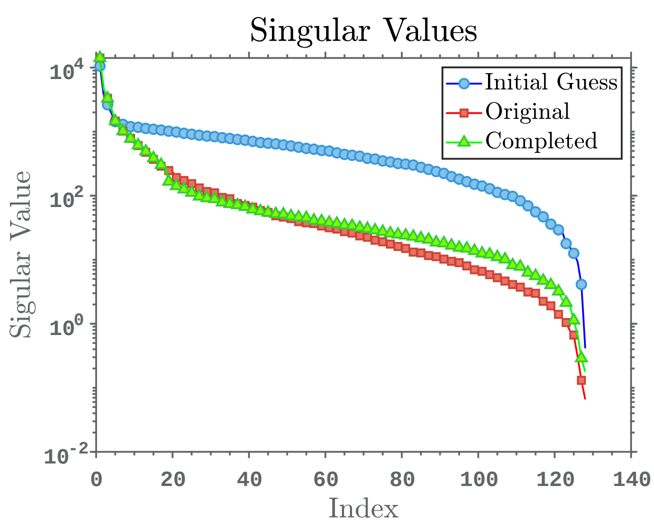


Figure 1.8: Singular values of Figure 1.5, Figure 1.7a, and Figure 1.7b

1.4 Alternating Minimization

The alternating minimization method is an empirically successful method for finding a low-rank completion of M_Ω. Moreover, it formed a critical component in the winning entry of the Netflix problem (Jain et al., 2013). Given a partially known matrix M_Ω, the objective of alternating minimization is to find a completed matrix in bilinear form $M = LR^\top$ with L being $m \times r$ and R being $n \times r$ such that the entries of $P_\Omega(M) = M_\Omega$. Such an M is found by alternating between optimizing L and R. In particular, the non-convex problem to solve is

$$\min_{L,R} \frac{1}{2} \left\| P_\Omega(LR^\top) - M_\Omega \right\|^2 .$$

The alternating minimization algorithm may be described in Algorithm 2.

Algorithm 2: Alternating Minimization (Jain et al., 2013)

Input: partially known matrix M_Ω, initial guess $R_0 \in M_{n\times r}$, stopping criterion

Result: $X_k = L_k R_k^\top$ an approximation of a rank r completion of M_Ω

for $k = 1, \ldots$ **do**

$$L_k = \underset{L \in M_{n\times r}}{\arg\min} \left\| P_\Omega(LR_{k-1}^\top) - M_\Omega \right\|^2;$$

$$R_k = \underset{R \in M_{m\times r}}{\arg\min} \left\| P_\Omega(L_k R^\top) - M_\Omega \right\|^2;$$

The minimization at each step may be solved with the method of least squares.

1.5 Orthogonal Rank-One Matrix Pursuit

Recall that we may express a matrix X as a weighted sum of rank one matrices M_i such that $\|M_i\| = 1$ for all i. That is, we may write X as

$$X = M(\theta) = \sum_i \theta_i M_i.$$

Here θ is the vector of weights in the sum. One way to calculate M_i and θ_i is with the singular value decomposition of X, by setting θ equal to the vector of singular values, and $M_i = u_i v_i^\top$ where u_i and v_i are the ith left and right-singular vectors of X respectively.

Note that the minimum value of $\|\theta\|_0$ over all choices of θ is equal to the rank of X, where $\|\theta\|_0$ is equal to the number of non-zero elements in the vector θ. Therefore, we may formulate the matrix completion problem as finding a solution to the minimization

$$\min_\theta \|P_\Omega(M(\theta)) - M_\Omega\|$$
$$\textit{s.t. } \|\theta\|_0 \leq r.$$

The goal of the orthogonal rank-one matrix pursuit algorithm is to find proper basis matrices M_i, and weights θ_i. We do so by alternating between computing the rank one basis matrices M_i and the weights θ accordingly. In particular, suppose we have computed $M_1, \ldots, M_{k-1}$ and weights θ^{k-1} on the $(k-1)$th step. To compute M_k, we first compute the regression residual

$$R_k = M_\Omega - \sum_{i=1}^{k-1} \theta_i M_i.$$

Because it is desired that M_k is rank one with unit Euclidean norm, we may search for M_k in the form $M_k = uv^\top$ for unit vectors u and v. We then calculate u and v as the solution to

$$\max_{u,v}\{u^\top R_k v \mid \|u\| = \|v\| = 1\}.$$

This minimization problem has optimal solution equal to the first left and right singular vectors of R_k. After we have computed $M_k = uv^\top$, we then calculate θ^k as the solution to the minimization problem

$$\min_\theta \left\| \sum_{i=1}^{k} \theta_i P_\Omega(M_i) - M_\Omega \right\| \tag{1.1}$$

which can be computed with least squares. In particular, let $m_\Omega = \text{vec}(M_\Omega)$, and $m_i = \text{vec}(P_\Omega(M_i))$ be the vectorization of M_Ω and $P_\Omega(M_i)$ respectively. Let $W_k = [m_1 \cdots m_k]$ be the vectors m_i assembled into a matrix for $i = 1, \ldots, k$. Then

$$\theta^k = (W_k^\top W_k)^{-1} W_k^\top m_\Omega$$

is the solution to Minimization (1.1).

In summary, the orthogonal rank-one matrix pursuit algorithm may be stated as follows.

Algorithm 3: Orthogonal Rank-One Matrix Pursuit (Wang et al., 2015)

Input: partially known matrix M_Ω, initial guess X_0, stopping criterion
Initialize: $m_\Omega = \text{vec}(M_\Omega)$
Result: X_k an approximation of a rank r completion of M_Ω
for $k = 1, \ldots$ **do**
  $R_k = M_\Omega - X_{k-1}$;
  Find the top left- and right-singular vectors u_k and v_k of R_k;
  $M_k = u_k v_k^\top$, $m_k = \text{vec}(M_k)$, and $W_k = [m_1 \cdots m_k]$;
  $\theta^k = (W_k^\top W_k)^{-1} W_k^\top m_\Omega$;
  $X_k = \sum_{i=1}^k \theta_i^k M_i$;

The orthogonal rank-one matrix pursuit algorithm converges at a linear rate (Wang et al., 2015).

1.6 Singular Value Thresholding

In general, the problem of finding a minimum rank completion of M_Ω is difficult because the rank function is non-convex. Moreover, the space of matrices with rank at most r, $\overline{\mathcal{M}_r}$, is low-dimensional, while real life data often has random noise. If we assume that the random noise is sampled from a continuous density, the probability that data with noise will belong to $\overline{\mathcal{M}_r}$ is zero.

Instead of solving $\min_X \text{rank}(X)$ such that $X_{ij} = M_{ij}$ for $(i, j) \in \Omega$, we may opt to solve a relaxation of the problem by replacing the rank function with the nuclear norm. In other words, we opt to solve the convex minimization $\min_X \|X\|_*$ such that $X_{ij} = M_{ij}$ for $(i, j) \in \Omega$. The *nuclear norm*

$\|X\|_*$ is defined as the sum of the singular values of the matrix X. That is, if $\sigma(X)$ is the vector of singular values of X and $\sigma_i(X)$ are the singular values, we have

$$\|X\|_* = \sum_i \sigma_i(X).$$

This convex relaxation is analogous to the ℓ^1 convex relaxation of the ℓ^0 norm for sparse signal recovery. In fact they are directly related, as we have

$$\begin{aligned} \operatorname{rank}(X) &= \|\sigma(X)\|_0 \\ \|X\|_* &= \|\sigma(X)\|_1 . \end{aligned}$$

where given a vector x, $\|x\|_0$ is the number of non-zero values in x, and $\|x\|_1$ is the sum of the magnitude of the entries in x. To understand why this relaxation is chosen, we introduce the definition of the convex envelope of a function.

Definition 13. *Given a convex domain C, the* convex envelope *of a function $f : C \to \mathbb{R}$ is the largest convex function g such that $g(x) \leq f(x)$.*

For example, the ℓ^1 norm is used as a convex relaxation of the ℓ^0 norm because it is the convex envelope of the ℓ^0 norm on the unit ball. Similarly, the nuclear norm is the convex envelope of the rank function on the unit ball with respect to the spectral norm. That is, on the domain $B = \{X \in M_{m\times n} \mid \sigma_1(X) \leq 1\}$.

Theorem 14. *(Fazel, 2002) On the unit ball $B = \{X \in M_{m\times n} \mid \sigma_1(X) \leq 1\}$, the convex envelope of the* rank *function is the nuclear norm function $\|\cdot\|_*$.*

Proof. First, recall that norms are convex functions, so the nuclear norm is convex. For $X \in B$, we have $\sigma_1(X) \leq 1$. Because $\sigma_1(X)$ is the largest singular value of X, we have $\sigma_i(X) \leq 1$ for all i. Let $r = \operatorname{rank}(X)$. Then r is the number of non-zero singular values of X. Therefore, we have

$$\|X\|_* = \sum_{i=1}^{r} \sigma_i(X) \leq \sum_{i=1}^{r} 1 = r,$$

so $\|X\|_* \leq \mathrm{rank}(X)$ for all $X \in B$. Moreover, it is shown in (Fazel, 2002) that the nuclear norm is the tightest convex lower bound. □

The rank minimization problem for matrix completion problem may then be approximated as the nuclear norm minimization problem

$$\begin{aligned}&\min_X \|X\|_* \\ &\textit{s.t. } P_\Omega(X) = M_\Omega.\end{aligned} \tag{1.2}$$

Singular value thresholding is an algorithm for approximately solving the convex minimization of the nuclear norm. Instead of directly minimizing the nuclear norm, we solve

$$\begin{aligned}&\min_X \tau \|X\|_* + \frac{1}{2}\|X\|^2 \\ &\textit{s.t. } P_\Omega(X) = M_\Omega\end{aligned} \tag{1.3}$$

which is easier to calculate. For large values of τ, the term $\tau \|X\|_*$ dominates $\frac{1}{2}\|X\|^2$, and so the solution is approximately equal to the solution to Minimization (1.2).

We start with defining the *singular value shrinkage operator* D_τ. Let $X = U\Sigma V^*$ be the singular value decomposition of X. Let $D_\tau(\Sigma)$ be equal to the diagonal matrix with ith diagonal entry equal to $\max(\sigma_i - \tau, 0)$. Then for $\tau \geq 0$, the singular value shrinkage operator D_τ is defined as $D_\tau(X) = UD_\tau(\Sigma)V^*$.

Theorem 15. *(Cai et al., 2010) For $\tau \geq 0$, the singular value shrinkage operator satisfies the minimization*

$$D_\tau(Y) = \arg\min_X \{\frac{1}{2}\|X - Y\|^2 + \tau\|X\|_*\}$$

Now in terms of the shrinkage operator, we present the singular value thresholding algorithm.

Algorithm 4: Singular Value Thresholding (Cai et al., 2010)

Input: partially known matrix M_Ω, sequence of step sizes $\{h_k\}$, $\tau \geq 0$, initial guess Y_0, stopping criterion

Result: X_k an approximation of a rank r completion of M_Ω

for $k = 1, \ldots$ **do**

  $X_k = D_\tau(Y_{k-1})$;

  $Y_k = Y_{k-1} + h_k(M_\Omega - P_\Omega(X_k))$;

It can be shown that the sequence X_k converges to the solution to Minimization (1.3), and for large values of τ, X_k approximates the solution to Minimization (1.2). Moreover, the matrices in the sequence $\{X_k\}$ empirically have low-rank (Cai et al., 2010).

1.7 Mask Equivalence by Permutation

Given an $m \times n$ partially known matrix M_Ω, it may be useful to permute the rows and columns or take the transpose to simplify the structure of the known and unknown entries. This does not change the number of rank r completions of M_Ω because row permutations, column permutations, and transposes are bijections under which the rank is invariant. More specifically, if π is a composition of transposes, row permutations, and column permutations, and M is a rank r completion of M_Ω, then $\pi(M)$ is a rank r completion of the partially known matrix $\pi(M_\Omega)$.

Recall that $B(\Omega)$ is the $m \times n$ binary matrix where the entry in position (i, j) is equal to one if $(i, j) \in \Omega$ and is equal to zero otherwise. We consider two masks $B(\Omega_1)$ and $B(\Omega_2)$ equivalent if we may obtain $B(\Omega_1)$ from $B(\Omega_2)$ by permuting rows and columns. How can we tell if two masks are equivalent? An initial attempt may be to count the number of ones in the rows and columns of $B(\Omega_1)$ and $B(\Omega_2)$. If $B(\Omega_1)$ and $B(\Omega_2)$ are equivalent, the number of ones in the rows and columns of $B(\Omega_1)$ and $B(\Omega_2)$ must be equal up to order respectively. This is a necessary, but not a sufficient, condition for two masks to be equivalent.

Example 16. *Consider the masks*

$$B(\Omega_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \qquad B(\Omega_2) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

The number of ones in the rows and columns are each $(1, 2, 2)$ *for both* $B(\Omega_1)$ *and* $B(\Omega_2)$*, so they may or may not be equivalent.*

Let $B_{m\times n}$ denote the set of $m \times n$ binary matrices. Because row and column permutations commute, there is a group action of $G = S_m \times S_n$ on $B_{m\times n}$, where S_n is the symmetric group over n symbols. In particular, for $(\alpha, \beta) \in S_m \times S_n$ and $X \in B_{m\times n}$,

$$(\alpha, \beta) \cdot X = Y$$

where Y is the matrix X with rows permuted by α and columns permuted by β.

Definition 17. *An element $g \in G$* stabilizes *an element $X \in B_{m\times n}$ if $g \cdot X = X$. The* stabilizer *of X is the set of group elements that stabilize X which we denote*

$$\mathrm{Stab}_G(X) = \{g \in G \mid g \cdot X = X\}.$$

Note that $B(\Omega_1)$ from Example 16 has a non-trivial stabilizer under the action of $S_3 \times S_3$ by permutation of rows and columns. In particular, swapping rows or columns two and three does not change $B(\Omega_1)$. On the other hand, $B(\Omega_2)$ is not fixed by swapping any two rows or columns, so it has a trivial stabilizer. Therefore, $B(\Omega_1)$ and $B(\Omega_2)$ are not equivalent.

In general, there are many equivalent masks.

Theorem 18. *The number of $m \times n$ masks up to equivalence is at least $\dfrac{2^{mn}}{m!n!}$.*

Proof. First note that $|B_{m\times n}| = 2^{mn}$. Let $G = S_m \times S_n$, then $|G| = m!n!$. Also G acts on $B_{m\times n}$, and $|B_{m\times n}/G|$ is the exact number of masks up to permutation of rows and columns.

By the orbit-counting theorem (Burnside, 1911), we have

$$|B_{m\times n}/G| = \frac{1}{|G|} \sum_{X\in B_{m\times n}} |\mathrm{Stab}_G(X)|\,.$$

Because each matrix is fixed by the identity permutation, we have $|\mathrm{Stab}_G(X)| \geq 1$ for all X, and so

$$|B_{m\times n}/G| \geq \frac{1}{|G|} \sum_{X\in B_{m\times n}} 1 = \frac{|B_{m\times n}|}{|G|} = \frac{2^{mn}}{m!n!}.$$

□

Given a partially known matrix, it may be useful to permute the rows and columns into more desirable positions. For example, if we know how to complete a partially known matrix with mask Ω_1, then it is likely that a matrix with mask Ω_2 will be completable if the rows and columns of Ω_2 may be permuted into Ω_1.

1.8 Finite Completability

A partially known matrix M_Ω may have finitely many, infinitely many, or no rank at most r completions. It is desirable to know if M_Ω has finitely many rank r completions or a unique rank r completion. If M_Ω has finitely many rank r completions, then knowledge of not many more entries of M_Ω will result in M_Ω having a unique completion.

Definition 19. *A partially known matrix M_Ω is called* finitely completable in rank r *if it has finitely many rank r completions. That is, if $0 < \left|\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r\right| < \infty$. M_Ω is called* uniquely completable in rank r *if it has a unique rank r completion, that is if $\left|\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r\right| = 1$.*

We discuss some necessary conditions for a partially known matrix M_Ω to have finitely many or a unique rank r completion.

Theorem 20. *Having at least r known entries per row and r known entries per column is a necessary condition for an $m \times n$ partially known matrix M_Ω to be finitely completable in rank r.*

Proof. Let Ω be a set of known positions such that there is a row or column of $B(\Omega)$ with fewer than r known entries. By transpose and permutation, suppose the last column of M_Ω has k known entries which is strictly fewer than r. Without loss of generality, assume the first $n-1$ columns are entirely known. Again by permuting the rows, let $M_\Omega = \left[\begin{smallmatrix} A & C \\ B & \square \end{smallmatrix}\right]$, where A and B are completely known, C is a $k \times 1$ block of known entries with $k < r$, and $\square$ is an $(m-k) \times 1$ block of unknown entries. It suffices to show that M_Ω is not finitely completable in rank r.

First note that if $\mathrm{rank}([\begin{smallmatrix} A \\ B \end{smallmatrix}]) > r$, then any completion of M_Ω will have rank greater than r, so there will be no completions of M_Ω in $\overline{\mathcal{M}}_r$. If $\mathrm{rank}([\begin{smallmatrix} A \\ B \end{smallmatrix}]) < r$, then any completion of M_Ω will have rank less than or equal to r, so M_Ω would have infinitely many rank r completions.

We are left with the case $\mathrm{rank}([\begin{smallmatrix} A \\ B \end{smallmatrix}]) = r$. Let $s = \mathrm{rank}(A)$. Note that $s \leq k$ Because A is a $k \times (n-1)$ matrix. Suppose C is not in the column space of A. In other words, suppose $\mathrm{rank}([\begin{smallmatrix} A & C \end{smallmatrix}]) = s+1$. Then any completion $X \in \mathcal{A}_\Omega$ will have rank $r+1$, and so there will be no completion of M_Ω in $\overline{\mathcal{M}}_r$. So C must be in the column space of A. In other words, $\mathrm{rank}([\begin{smallmatrix} A & C \end{smallmatrix}]) = s$.

Because $\mathrm{rank}([\begin{smallmatrix} A \\ B \end{smallmatrix}]) = r$, there exists an $r \times (n-1)$ rank r submatrix $\left[\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right]$. We may choose A' such that it consists of s linearly independent rows of A. Then we must choose the remaining $r-s$ linearly independent rows B' from B because the remaining rows of A are in the row space of A'.

Augmenting $\left[\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right]$ with the corresponding rows from $[\begin{smallmatrix} C \\ \square \end{smallmatrix}]$, we get an $r \times n$ submatrix of M_Ω of the form $M'_\Omega = \left[\begin{smallmatrix} A' & C' \\ B' & \square \end{smallmatrix}\right]$, where $\square$ is an $(r-s) \times 1$ block of unknown entries. Any completion $M' = \left[\begin{smallmatrix} A' & C' \\ B' & D' \end{smallmatrix}\right]$ of M'_Ω will be rank r because the submatrix $\left[\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right]$ is rank r. Moreover, there are $r-s$ degrees of freedom, which is greater than zero because $r > k \geq s$.

Let $B = \left[\begin{smallmatrix} B' \\ B'' \end{smallmatrix}\right]$, where B'' is the $(m-k-r+s) \times (n-1)$ submatrix of B consisting of the rows in B that are not in B'. Similarly, let $D = [\begin{smallmatrix} D' \\ \square \end{smallmatrix}]$, where $\square$ are the remaining $(m-k-r+s)$ unknown entries of M_Ω. Because the rows of $\left[\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right]$ form a basis for the row space of $[\begin{smallmatrix} A \\ B \end{smallmatrix}]$, there exists a unique $r \times (m-k-r+s)$ matrix X such that $\left[\begin{smallmatrix} A' \\ B' \end{smallmatrix}\right]^\top X = [\begin{smallmatrix} B'' \end{smallmatrix}]^\top$. In particular, $X = \left[\begin{smallmatrix} A'(A')^\top & A'(B')^\top \\ B'(A')^\top & B'(B')^\top \end{smallmatrix}\right]^{-1} \left[\begin{smallmatrix} A'(B'')^\top \\ B'(B'')^\top \end{smallmatrix}\right]$. So we must have $\left[\begin{smallmatrix} A' & C' \\ B' & D' \end{smallmatrix}\right]^\top X = [\begin{smallmatrix} B'' & \square \end{smallmatrix}]^\top$, which implies the remaining unknown entries are equal to $\left[\begin{smallmatrix} C' \\ D' \end{smallmatrix}\right]^\top X$. So there exists a unique D such that $\mathrm{rank}([\begin{smallmatrix} A' & C' \\ B & D \end{smallmatrix}]) = r$.

Moreover, because the rank of a matrix is at least as great as the rank of any submatrix, we have

$$s = \mathrm{rank}(A') \leq \mathrm{rank}([\begin{smallmatrix} A' & C' \end{smallmatrix}]) \leq \mathrm{rank}([\begin{smallmatrix} A & C \end{smallmatrix}]) = s.$$

So $\mathrm{rank}([\begin{smallmatrix} A' & C' \end{smallmatrix}]) = s$, which implies that the rows of $[\begin{smallmatrix} A' & C' \end{smallmatrix}]$ span the row space of $[\begin{smallmatrix} A & C \end{smallmatrix}]$. Therefore $\mathrm{rank}([\begin{smallmatrix} A & C \\ B & D \end{smallmatrix}]) = \mathrm{rank}([\begin{smallmatrix} A' & C' \\ B & D \end{smallmatrix}]) = r$.

Thus any completion $M' = \left[\begin{smallmatrix} A' & C' \\ B' & D' \end{smallmatrix}\right]$ of M'_Ω extends to a unique rank r completion $M = \left[\begin{smallmatrix} A & C \\ B & D \end{smallmatrix}\right]$ of M_Ω. So $\dim(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = r - s > 0$, and so M_Ω has infinitely many rank r completions.

This exhausts all cases of M_Ω, so M_Ω is not finitely completable in rank r. □

It was noted in (Candès & Tao, 2010) that at least $(m+n)r - r^2$ entries must be observed for M_Ω to be finitely completable in rank r. We give a precise proof.

Theorem 21. *Given a partially known $m \times n$ matrix M_Ω, Ω must contain at least $(m+n)r - r^2$ known positions for M_Ω to be finitely completable.*

Proof. From proposition I.7.1 in (Hartshorne, 2013), if V and W are irreducible algebraic varieties in d dimensional affine space, then we have the inequality $\dim(V) + \dim(W) \leq \dim(V \cap W) + d$.

Note that $\dim(\overline{\mathcal{M}}_r) = (m+n)r - r^2$, $\dim(\mathcal{A}_\Omega) = mn - |\Omega|$, and $\dim(M_{m\times n}) = mn$. Suppose M_Ω is finitely completable in rank r, then $\dim(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = 0$. Applying the inequality, we must have $(m+n)r - r^2 + mn - |\Omega| \leq mn$, which implies $(m+n)r - r^2 \leq |\Omega|$. □

Given M_Ω, it may not be known a priori which rank r should be chosen to complete M_Ω. If r is chosen too small, then $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ will be empty, and if r is chosen too large, then M_Ω will have infinitely many completions. We give a method of deciding such a rank r.

Suppose we are given Ω, but we do not know r. Let $p = |\Omega|$. Then if M_Ω is finitely completable in rank r, we must have $p \geq (m+n)r - r^2$ from Theorem 21. This implies

$$r \leq \frac{m+n-\sqrt{(m+n)^2 - 4p}}{2}.$$

So a good guess for a rank r may be $r = \left\lfloor \frac{m+n-\sqrt{(m+n)^2-4p}}{2} \right\rfloor$. It is worth noting that for many methods, greater than $(m+n)r - r^2$ observed entries are required for convergence, in which case r may need to be chosen smaller.

It is useful to define the function $\Phi_\Omega : \overline{\mathcal{M}}_r \to M_{m\times n}$ as the restriction of P_Ω to $\overline{\mathcal{M}}_r$. In other words, $\Phi_\Omega(X)$ is the projection of a matrix $X \in \overline{\mathcal{M}}_r$ obtained by setting entries in positions not in Ω equal to zero. Then given a partially known matrix M_Ω, we have $\Phi_\Omega^{-1}(M_\Omega) = \mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$. In other words, $\Phi_\Omega^{-1}(M_\Omega)$ is the space of rank at most r completions of M_Ω.

We now focus on the set of matrices X such that given Ω, the preimage $\Phi_\Omega^{-1}(\Phi_\Omega(X))$ is a zero dimensional set. In other words, the set of $X \in \overline{\mathcal{M}}_r$ such that $\Phi_\Omega(X)$ is finitely completable in rank r. Define the set $\chi_\Omega \subset \overline{\mathcal{M}}_r$ as

$$\chi_\Omega = \left\{X \in \overline{\mathcal{M}}_r \mid \dim(\Phi_\Omega^{-1}(\Phi_\Omega(X))) = 0\right\}.$$

Theorem 22. *For any $X \in \chi_\Omega$, $\operatorname{rank}(X) = r$. That is, $\chi_\Omega \subset \mathcal{M}_r$.*

Proof. By permuting rows and columns, suppose the position $(1,1) \notin \Omega$. Consider $X \in \chi_\Omega$ such that

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Suppose $\operatorname{rank}(X) < r$. Then the matrix

$$Y(t) = \begin{bmatrix} t & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

has rank at most r for any t. So $Y(t) \in \overline{\mathcal{M}}_r$ for any t. Moreover $Y(t) \in \Phi_\Omega^{-1}(\Phi_\Omega(X))$ for any t because $\Phi_\Omega(X) = \Phi_\Omega(Y(t))$. However, this implies that $\dim(\Phi_\Omega^{-1}(\Phi_\Omega(X))) > 0$, contradicting the assumption that $X \in \chi_\Omega$. Therefore, we must have $\operatorname{rank}(X) = r$. □

By Theorem 22, if M_Ω is finitely completable in rank r, then each rank at most r completion of M_Ω has rank equal to r. To obtain an upper bound on the number of rank r completions, we introduce the *degree* of a variety.

Definition 23. *The* degree *of an algebraic variety V of dimension k is the number of intersection points of V with k hyperplanes in general position.*

For example, the degree of the algebraic variety $\overline{\mathcal{M}}_r$ is known.

Theorem 24. *(Harris & Tu, 1984) The degree of the algebraic variety $\overline{\mathcal{M}}_r$ is*

$$\deg(\overline{\mathcal{M}}_r) = \prod_{i=0}^{n-r-1} \frac{\binom{m+i}{n-1-i}}{\binom{m-r+i}{n-r-i}}.$$

Recall a generalized version of Bézout's Theorem.

Theorem 25. *(Fulton, 2016) Let $V_1, \ldots, V_k$ be irreducible algebraic varieties, and let $Z_1, \ldots, Z_N$ be the irreducible components of $V_1 \cap \cdots \cap V_k$. Then*

$$\sum_{i=1}^{N} \deg(Z_i) \leq \prod_{j=1}^{k} \deg(V_j).$$

Now we are ready to present an upper bound on the number of rank r completions of M_Ω.

Theorem 26. *(Lai & Varghese, 2017) Given a partially known $m \times n$ matrix M_Ω, suppose M_Ω is finitely completable in rank r. Let $N = \left|\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r\right|$ be the number of rank r completions of M_Ω, then $N \leq \deg(\overline{\mathcal{M}}_r)$.*

Proof. Given M_Ω, recall that $\mathcal{A}_\Omega = P_\Omega^{-1}(M_\Omega)$ is the algebraic variety of completions of M_Ω. Note that $\mathcal{A}_\Omega$ is a linear variety, and so it is irreducible and $\deg(\mathcal{A}_\Omega) = 1$.

Now suppose there are finitely many points $Z_1, \ldots, Z_N \in \mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$. That is, there are N rank r completions of M_Ω. Because $\deg(Z_i) = 1$ for all i, and both $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$ are irreducible, then by Theorem 24 and Theorem 25 we have

$$N = \sum_{i=1}^{N} \deg(Z_i) \leq \deg(\mathcal{A}_\Omega) \deg(\overline{\mathcal{M}}_r) = \deg(\overline{\mathcal{M}}_r)$$

□

In general, $\deg(\overline{\mathcal{M}}_r)$ is very large, and is larger than the exact number of rank at most r completions. One reason for this may be that given M_Ω, the hyperplanes $H_{ij} = \{X \in M_{m\times n} \mid X_{ij} = M_{ij}\}$ such that $\cap_{(i,j)\in\Omega} H_{ij} = \mathcal{A}_\Omega$ may not be in general position. In addition, some intersection points may be at infinity, or the intersection points may have multiplicity.

It is often desirable to have a unique rank r completion rather than finitely many. If Ω has the right structure, then a generic $M \in \overline{\mathcal{M}}_r$ will be the unique rank r completion of the partially known matrix $M_\Omega = \Phi_\Omega(M)$. We introduce the following definition.

Definition 27. *A mask Ω is called* finitely completable in rank r *or* uniquely completable in rank r *if, for a generic $M \in \overline{\mathcal{M}}_r$, the partially known matrix $M_\Omega = \Phi_\Omega(M)$ is finitely completable or uniquely completable respectively.*

We give a class of Ω which are uniquely completable in rank r. First we introduce the following definition and theorem.

Definition 28. *A mask Ω is called* completable entry by entry in r *if we may find an $(r+1)\times(r+1)$ submatrix of $B(\Omega)$ with exactly one entry equal to zero, replace that zero with a one, and repeat until all entries are equal to one.*

Theorem 29. *If a mask Ω is completable entry by entry in r, then it is uniquely completable in r.*

To prove Theorem 29, we introduce the following lemmas. First, we recall a lemma from exercise I.1.8 in (Hartshorne, 2013).

Lemma 30. *(Hartshorne, 2013) Let V be an irreducible algebraic variety in affine space $\mathbb{A}^d$. Let H be a hypersurface in $\mathbb{A}^d$, meaning $\dim(H) = d-1$. Then either $\dim(V\cap H) = \dim(V) - 1$, or $V \subset H$.*

We use Lemma 30 to prove Lemma 31. First, we introduce the following notation. Given an $m \times n$ matrix X, and sets of indices $I \subset [m]$ and $J \subset [n]$, define $X_{I,J}$ as the submatrix of X with entries in positions in $I \times J$. Note that $X_{I,J}$ is only defined up to permutations of rows and columns because I and J are not ordered. However, the absolute value of the determinant of $X_{I,J}$ is well-defined because swapping rows or columns only changes the sign of the determinant.

Lemma 31. *Let H be the algebraic variety consisting of $m \times n$ matrices that have at least one vanishing $r \times r$ minor. Then* $\dim(\overline{\mathcal{M}}_r \cap H) < \dim(\overline{\mathcal{M}}_r)$.

Proof. Given $I \subset [m]$, $J \subset [n]$, such that $|I| = |J| = r$, let

$$H_{I,J} = \{X \in M_{m \times n} \mid \det(X_{I,J}) = 0\}.$$

Then $H_{I,J}$ is a hypersurface because it is the zero set of one polynomial equation, and

$$H = \bigcup_{|I|=|J|=r} H_{I,J}$$

is the algebraic variety of $m \times n$ matrices with at least one vanishing $r \times r$ minor. Then

$$\overline{\mathcal{M}}_r \cap H = \bigcup_{|I|=|J|=r} (\overline{\mathcal{M}}_r \cap H_{I,J})$$

So it suffices to show that $\dim(\overline{\mathcal{M}}_r \cap H_{I,J}) < \dim(\overline{\mathcal{M}}_r)$ for all $|I| = |J| = r$. Note that $\overline{\mathcal{M}}_r$ is not contained in $H_{I,J}$ because there are $m \times n$ rank r matrices such that the $r \times r$ submatrix with positions in $I \times J$ is nonsingular. For example, the matrix consisting of the $r \times r$ identity matrix in position $I \times J$ and all other entries equal to zero. So by Lemma 30,

$$\dim(\overline{\mathcal{M}}_r \cap H_{I,J}) = \dim(\overline{\mathcal{M}}_r) - 1 < \dim(\overline{\mathcal{M}}_r)$$

for all $|I| = |J| = r$ because $\overline{\mathcal{M}}_r$ is irreducible and not contained in $H_{I,J}$. □

We now define the entry by entry matrix completion algorithm, Algorithm 5, which we then use to prove Theorem 29.

Proof. Given a mask Ω, suppose Ω is completable entry by entry in r. Then given $M \in \overline{\mathcal{M}}_r$, input $M_\Omega = \Phi_\Omega(M)$ into Algorithm 5. At each iteration we may always find a submatirx with exactly one unknown entry because Ω is completable entry by entry in r. The only way Algorithm 5 may output a

Algorithm 5: Complete Entry by Entry

Input: mask Ω_0, partially known matrix M_{Ω_0}, rank r
Result: mask Ω_k such that $|\Omega_k| \geq |\Omega_0|$, partially known matrix M_{Ω_k} with same rank at most r completions as M_{Ω_0}

for $k = 0, 1, \ldots$ **do**
  find an $(r+1) \times (r+1)$ submatrix of M_{Ω_k} of the form $X_k = \left[\begin{smallmatrix} A_k & b_k \\ c_k & x_k \end{smallmatrix}\right]$ where A_k is $r \times r$ nonsingular and x_k is the only unknown element in some position (i_k, j_k) in X_k;
  if *no such submatrix exists* **then**
    return Ω_k and M_{Ω_k};
  else
    solve $\det(X_k) = 0$ for x_k getting $x_k = \frac{a_k}{\det(A_k)}$ for a constant a_k;
    $M_{\Omega_{k+1}}$ equals M_{Ω_k} with the entry in position (i_k, j_k) equal to $\frac{a_k}{\det(A_k)}$;
    $\Omega_{k+1} = \Omega_k \cup \{(i_k, j_k)\}$;

matrix that is not fully completed is if there is at least one $r \times r$ singular submatrix in M. However, by Lemma 31, the set of matrices with at least one singular $r \times r$ submatrix in $\overline{\mathcal{M}}_r$ has dimension strictly smaller than the dimension of $\overline{\mathcal{M}}_r$, so a matrix with at least one singular $r \times r$ submatrix is not generic in $\overline{\mathcal{M}}_r$. So Algorithm 5 will output a fully completed matrix X for a generic $M \in \overline{\mathcal{M}}_r$. Moreover, because M satisfies the equations used to complete M_Ω, each of which had a unique solution, we must have $X = M$. □

More generally, given any mask Ω and any $M \in \overline{\mathcal{M}}_r$, let $M_{\Omega'}$ be the output of Algorithm 5 with input $M_\Omega = \Phi_\Omega(M)$. Then all completed entries will be equal to the corresponding entry in M. Again, this is because each equation used to complete M_Ω had a unique solution, and M satisfies those equations. So the entries completed by Algorithm 5 must be equal to the corresponding entry in every rank at most r completion $X \in \Phi_\Omega^{-1}(M_\Omega)$. So the space of rank at most r completions of M_Ω must be equal to the space of rank at most r completions of $M_{\Omega'}$. That is, we have $\Phi_{\Omega'}^{-1}(M_{\Omega'}) = \Phi_\Omega^{-1}(M_\Omega)$.

Example 32. *A special class of partially known matrices which is are completable entry by entry, and thus have a unique rank r completion, are matrices of the form* $M_\Omega = \left[\begin{smallmatrix} A & B \\ C & \square \end{smallmatrix}\right]$. *Here* A *is* $r \times r$ *and nonsingular,* B *is* $r \times (n-r)$, C *is* $(m-r) \times r$, *and* A, B, *and* C *consist of the known entries of* M_Ω. *In fact,* M_Ω *may be completed all at once to the unique rank r completion* $M = \left[\begin{smallmatrix} A & B \\ C & CA^{-1}B \end{smallmatrix}\right]$. *Note that* M_Ω *has* $(m+n)r - r^2$

known entries, which is the minimum number of known entries such that M_Ω may have a unique completion by Theorem 21.

1.9 Reduction of Generating Set

Recall that $\overline{\mathcal{M}}_r$ is the algebraic variety generated by all $(r+1) \times (r+1)$ minors. In other words,

$$\overline{\mathcal{M}}_r = \{X \in M_{m\times n} \mid \det(X_{I,J}) = 0, |I| = |J| = r+1\}$$

where $I \subset [m]$ and $J \subset [n]$. Given a partially known matrix M_Ω with known entries $\{M_{ij}\}_{(i,j)\in\Omega}$, $\mathcal{A}_\Omega$ is generated by the polynomials $X_{ij} - M_{ij}$ for all $(i,j) \in \Omega$. In other words,

$$\mathcal{A}_\Omega = \{X \in M_{m\times n} \mid X_{ij} - M_{ij} = 0, (i,j) \in \Omega\}$$

Therefore, $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is generated by the union of these polynomials. However, this set of generators may be redundant in the sense that $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ may be generated by a proper subset of these polynomials.

In Example 32, because A is a full rank $r\times r$ known submatrix of M_Ω and B and C are fully known, we may complete M_Ω exclusively with equations of the form $\left|\begin{smallmatrix} A & b \\ c & x \end{smallmatrix}\right| = 0$. Here x is an unknown element, b is a column of B, and c is a row of C, getting $x = cA^{-1}b$. In other words, we don't need all $(r+1)\times(r+1)$ minors to find the rank r completion of M_Ω, we only need the $(r+1)\times(r+1)$ minors that contain A. What if B and C aren't fully known? What if A is smaller than $r \times r$? Can we still recover the rank r completions of M_Ω using only the minors that contain A under the assumption that A is full rank? We show that the answer is yes.

When M_Ω contains some known full rank submatrix A, we show that there are many redundant polynomial generators for $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ in the union of the generators for $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$. Moreover, M_Ω always contains such a submatrix, as one may always at least take A as a 1×1 known nonzero entry. If every known entry of M_Ω is equal to zero, then M_Ω has the unique rank zero completion equal to the zero matrix.

We start with the following lemma.

Lemma 33. *Let $V_{q,n} \subset M_{(r+1)\times n}$ be the algebraic variety generated by all $(r+1)\times(r+1)$ minors containing the first q columns where $q \leq r+1 \leq n$. In other words,*

$$V_{q,n} = \{X \in M_{(r+1)\times n} \mid \det(X_{[r+1],J}) = 0, [q] \subset J\}$$

where $|J| = r+1$. Let $W_{q,n} \subset M_{(r+1)\times n}$ be the algebraic variety consisting of matrices where the first q columns are linearly dependant. That is,

$$W_{q,n} = \{X \in M_{(r+1)\times n} \mid \textit{the first } q \textit{ columns of } X \textit{ are linearly dependant}\}.$$

Then

$$V_{q,n} = \overline{\mathcal{M}}_r \cup W_{q,n}.$$

Note that when $q = 1$, the only way a single column can be linearly dependant is if that column is the zero vector. Also note that the first q columns of a matrix $X \in W_{q,n}$ is a submatrix which has rank at most $q-1$. Moreover, X has $(r+1)(n-q)$ free parameters. Therefore,

$$W_{q,n} \cong \overline{\mathcal{M}}_{q-1} \times M_{(r+1)\times(n-q)}$$

where $\overline{\mathcal{M}}_{q-1} \subset M_{(r+1)\times q}$. Moreover, $W_{q,n}$ is irreducible because both $\overline{\mathcal{M}}_{q-1}$ and $M_{(r+1)\times(n-q)}$ are irreducible and the product of irreducible algebraic varieties is irreducible. Therefore, $V_{q,n} = \overline{\mathcal{M}}_r \cup W_{q,n}$ is the decomposition of $V_{q,n}$ into irreducible components.

We prove Lemma 33 by inducting on n and backwards inducting on q.

Proof. Note that $\overline{\mathcal{M}}_r \subset V_{q,n}$ because the set of polynomials which generate $\overline{\mathcal{M}}_r$ contain the set of polynomials which generate $V_{q,n}$. Also, $W_{q,n} \subset V_{q,n}$ because if $X \in W_{q,n}$, then the first q columns of X are linearly dependent, so every $(r+1) \times (r+1)$ minor containing the first q columns vanishes, so $X \in V_{q,n}$. Therefore, $\overline{\mathcal{M}}_r \cup W_{q,n} \subset V_{q,n}$.

For the reverse inclusion, first consider the case $n = r+1$, and $q \leq n$ is arbitrary. There is exactly one $(r+1) \times (r+1)$ minor which generates both $V_{q,r+1}$ and $\overline{\mathcal{M}}_r$, so $V_{q,r+1} = \overline{\mathcal{M}}_r$. Also, any $X \in W_{q,r+1}$ will be be singular for all $q \leq r+1$, so $W_{q,r+1} \subset \overline{\mathcal{M}}_r$, and so $V_{q,r+1} = \overline{\mathcal{M}}_r = \overline{\mathcal{M}}_r \cup W_{q,r+1}$.

Now consider the case $q = r+1$ and $n \geq q$ is arbitrary. Then both $V_{r+1,n}$ and $W_{r+1,n}$ are generated by the first $(r+1) \times (r+1)$ minor. That is, the polynomial $\det(X_{I,J})$ where $I = J = [r+1]$, so $V_{r+1,n} = W_{r+1,n}$. Moreover, $\overline{\mathcal{M}}_r \subset W_{r+1,n}$ because the polynomial that generates $W_{r+1,n}$ is an element of the set of polynomials which generate $\overline{\mathcal{M}}_r$, so $V_{r+1,n} = W_{r+1,n} = \overline{\mathcal{M}}_r \cup W_{r+1,n}$.

Now we assume that $V_{q,n-1} = \overline{\mathcal{M}}_r \cup W_{q,n-1}$, and $V_{q+1,n} = \overline{\mathcal{M}}_r \cup W_{q+1,n}$ for some n and q, and we seek to show $V_{q,n} = \overline{\mathcal{M}}_r \cup W_{q,n}$.

Consider some $X \in V_{q,n}$. If $\operatorname{rank}(X) \leq r$, then $X \in \overline{\mathcal{M}}_r$, so suppose $\operatorname{rank}(X) = r+1$. Let Y be the $(r+1) \times (n-1)$ submatrix obtained by deleting the $(q+1)$st column of X. Then because $X \in V_{q,n}$, we have $Y \in V_{q,n-1}$ because every minor containing the first q columns of Y vanishes. Moreover because $V_{q,n-1} = \overline{\mathcal{M}}_r \cup W_{q,n-1}$, then $Y \in \overline{\mathcal{M}}_r$ or $Y \in W_{q,n-1}$. In other words, $\operatorname{rank}(Y) \leq r$ or the first q columns of Y are linearly dependent.

If $Y \in W_{q,n-1}$ then $X \in W_{q,n}$ because the first q columns of Y and X are identical. So suppose $Y \in \overline{\mathcal{M}}_r$, then $\operatorname{rank}(Y) \leq r$. Because $\operatorname{rank}(X) = r+1$, we must have that $\operatorname{rank}(Y) = r$, and the $(q+1)$st column of X is linearly independent from the other columns. In particular, it is linearly independent from the first q columns.

Now note that $X \in V_{q+1,n}$ because $V_{q,n} \subset V_{q+1,n}$. By assumption, $V_{q+1,n} = \overline{\mathcal{M}}_r \cup W_{q+1,n}$, so $X \in \overline{\mathcal{M}}_r$ or $X \in W_{q+1,n}$. We have already assumed that X is not in $\overline{\mathcal{M}}_r$, so we must have $X \in W_{q+1,n}$, which means the first $q+1$ columns of X must be linearly dependent. However, the $(q+1)$st column is linearly independent from the first q columns, which means that the first q columns are linearly dependent, so $X \in W_{q,n}$. Therefore, $X \in \overline{\mathcal{M}}_r \cup W_{q,n}$, which implies $V_{q,n} \subset \overline{\mathcal{M}}_r \cup W_{q,n}$. So $V_{q,n} = \overline{\mathcal{M}}_r \cup W_{q,n}$. So by induction, $V_{q,n} = \overline{\mathcal{M}}_r \cup W_{q,n}$ for all $n \geq r+1$ and all $q \leq r+1$ □

More generally, given positions of any q columns, if V is the algebraic variety generated by the minors containing those q columns and W is the set of matrices such that columns in those positions are linearly dependant, then $V = \overline{\mathcal{M}}_r \cup W$. This follows simply by permuting the columns to the first q positions. We use the Lemma 33 to prove the following generalization.

Theorem 34. *Let $V_{p,q,m,n} \subset M_{m\times n}$ be the algebraic variety generated by all $(r+1)\times(r+1)$ minors containing all entries in positions $[p]\times[q]$. That is, the top-left $p \times q$ submatrix, with* $\max(p,q) \leq r+1 \leq \min(m,n)$. *In other words,*

$$V_{p,q,m,n} = \{X \in M_{m\times n} \mid \det(X_{I,J}) = 0, [p] \subset I, [q] \subset J\}$$

where $|I| = |J| = r+1$. Let $W_{p,q,m,n} \subset M_{m\times n}$ be the algebraic variety consisting of $m \times n$ matrices such that the top-left $p \times q$ submatrix is not full rank. That is,

$$W_{p,q,m,n} = \{X \in M_{m\times n} \mid \textit{the submatrix } X_{[p],[q]} \textit{ is not full rank}\}.$$

Then

$$\overline{\mathcal{M}}_r \subset V_{p,q,m,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,m,n}.$$

Proof. $\overline{\mathcal{M}}_r \subset V_{p,q,m,n}$ because the set of polynomials that generate $V_{p,q,m,n}$ are a subset of the set of polynomials that generate $\overline{\mathcal{M}}_r$.

To show $V_{p,q,m,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,m,n}$ we induct on m. By transposing, suppose $q \leq p$. Let $m = r+1$. Consider some $X \in V_{p,q,r+1,n}$, and let $A = X_{[p],[q]}$ be the top-left $p \times q$ submatrix of X. If $\operatorname{rank}(A) < q$, then A is not full rank, so $X \in W_{p,q,r+1,n}$.

Suppose $\operatorname{rank}(A) = q$. Note that $V_{p,q,r+1,n} = V_{q,n}$ as defined in Lemma 33, so $X \in \overline{\mathcal{M}}_r$ or the first q columns of X are linearly dependant. However, the first q columns of X are linearly independent because $\operatorname{rank}(A) = q$, so $X \in \overline{\mathcal{M}}_r$. So $V_{p,q,r+1,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,r+1,n}$

Now suppose $V_{p,q,m,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,m,n}$ for some $m \geq r+1$. Consider some $X \in V_{p,q,m+1,n}$. Let $A = X_{[p],[q]}$ be the top-left $p \times q$ submatrix of X. If $\operatorname{rank}(A) < q$, then $X \in W_{p,q,m+1,n}$.

Suppose $\operatorname{rank}(A) = q$. Consider the $m \times n$ submatrix Y of X which consists of X without the $(p+1)$st row, and exists because $m+1 \geq p+1$. Then $Y \in V_{p,q,m,n}$ because all $(r+1) \times (r+1)$ minors containing A vanish. By assumption, $V_{p,q,m,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,m,n}$, so $Y \in \overline{\mathcal{M}}_r$ or $Y \in W_{p,q,m,n}$. However, Y is not in $W_{p,q,m,n}$ because A is full rank, so $Y \in \overline{\mathcal{M}}_r$ which means $\operatorname{rank}(Y) \leq r$.

If $\operatorname{rank}(Y) < r$, then by returning the $(p+1)$st row of X back into Y, we have $\operatorname{rank}(X) \leq r$, which means $X \in \overline{\mathcal{M}}_r$. Now suppose $\operatorname{rank}(Y) = r$. Let Q be a submatrix of A consisting of q linearly independent rows of A, which exists because $\operatorname{rank}(A) = q$. Then the q corresponding rows in Y which contain Q are also linearly independent. Let R be an $r \times n$ submatrix of Y consisting of r linearly independent rows of Y which include the q linearly independent rows containing Q. Then the rows of R form a basis for the row space of Y.

Now consider the $m \times n$ submatrix Z of X obtained by removing one row which is not contained in R and is not the $(p+1)$st row of X. Such a row exists because $m+1 \geq r+2$.

Let B be the top-left $p \times q$ submatrix of Z. Then Z contains Q, so $\operatorname{rank}(B) = q$, which means B is full rank. Then like before, $Z \in V_{p,q,m,n}$, which means $Z \in \overline{\mathcal{M}}_r$ or $Z \in W_{p,q,m,n}$. Because B is full rank, we must have $Z \in \overline{\mathcal{M}}_r$, so $\operatorname{rank}(Z) \leq r$. Because Z contains R, we also have $\operatorname{rank}(Z) \geq r$. So $\operatorname{rank}(Z) = r$, and the rows of R form a basis for the row space of Z.

Therefore, the $(p+1)$st row of X which is contained in Z but not contained in R is contained in the row space of R, so the rows of R form a basis for the row space of X, which means $\operatorname{rank}(X) = r$, so $X \in \overline{\mathcal{M}}_r$.

We have shown that $X \in V_{p,q,m+1,n}$ implies $X \in \overline{\mathcal{M}}_r \cup W_{p,q,m+1,n}$, so $V_{p,q,m+1,n} \subset \overline{\mathcal{M}}_r \cup W_{p,p,m+1,n}$. So by induction, $V_{p,q,m,n} \subset \overline{\mathcal{M}}_r \cup W_{p,q,m,n}$ for any $m \geq r+1$. □

Like Lemma 33, Theorem 34 generalizes straightforwardly.

Theorem 35. *Given index sets $I \subset [m]$ and $J \subset [n]$, and let $p = |I|$ and $q = |J|$. Let $V \subset M_{m\times n}$ be the algebraic variety generated by all $(r+1) \times (r+1)$ minors which contain all entries in positions $I \times J$,* *where* $\max(p, q) \leq r + 1 \leq \min(m, n)$. *Let* $W \subset M_{m\times n}$ *be the algebraic variety consisting of* $m \times n$ *matrices such that the* $p \times q$ *submatrix with entries in positions* $I \times J$ *is not full rank. Then*

$$\overline{\mathcal{M}}_r \subset V \subset \overline{\mathcal{M}}_r \cup W.$$

Proof. This follows from Theorem 34 by permuting the rows in positions I to positions $[p]$, and the columns in positions J to positions $[q]$. □

Theorem 35 is important because if we are given a partially known matrix M_Ω, then M_Ω will contain some full rank known submatrix, which we may use to reduce the number of equations which generate $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$. More specifically, we have the following theorem.

Theorem 36. *Given* M_Ω*, find a full rank submatrix* A *consisting of known entries in some positions* $I \times J$. *Let* $|I| = p$ *and* $|J| = q$. *Let* V *be the zero set of all* $(r + 1) \times (r + 1)$ *minors containing the positions* $I \times J$ *for some* $r \geq \max(p, q)$. *Then*

$$\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r = \mathcal{A}_\Omega \cap V.$$

Proof. By Theorem 34, we have $\overline{\mathcal{M}}_r \subset V \subset \overline{\mathcal{M}}_r \cup W$ where W is the set of matrices such that for all $X \in W$, the submatrix $X_{I,J}$ is not full rank. Intersecting with $\mathcal{A}_\Omega$, we have

$$\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r \subset \mathcal{A}_\Omega \cap V \subset \mathcal{A}_\Omega \cap (\overline{\mathcal{M}}_r \cup W) = (\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) \cup (\mathcal{A}_\Omega \cap W)$$

Because $X_{I,J}$ is full rank for all $X \in \mathcal{A}_\Omega$ and $X_{I,J}$ is not full rank for all $X \in W$, $\mathcal{A}_\Omega \cap W$ is empty. So

$$\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r \subset \mathcal{A}_\Omega \cap V \subset \mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$$

which means $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r = \mathcal{A}_\Omega \cap V$. □

Theorem 36 means, given a partially known matrix M_Ω, if we find some full rank known submatrix A of M_Ω consisting of known entries, then $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is generated by all $(r+1)\times(r+1)$ minors containing A, along with the polynomials $X_{ij} - M_{ij}$ for all $(i,j) \in \Omega$.

There are $\binom{m}{r+1}\binom{n}{r+1}$ submatrices of size $(r+1)\times(r+1)$ in an $m \times n$ matrix up to permutations of rows and columns, which means there are $\binom{m}{r+1}\binom{n}{r+1}$ equations which generate $\overline{\mathcal{M}}_r$. Given M_Ω, there are $|\Omega|$ equations which generate $\mathcal{A}_\Omega$ which is equal to the number of known entries in M_Ω. In total, there are $\binom{m}{r+1}\binom{n}{r+1} + |\Omega|$ equations which generate $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ as the union of the generators of $\overline{\mathcal{M}}_r$ and $\mathcal{A}_\Omega$.

There are $\binom{m-p}{r+1-q}\binom{n-q}{r+1-q}$ submatrices of size $(r+1)\times(r+1)$ which contain a $p \times q$ matrix, so if one can find a fully known and full rank $p \times q$ submatrix of M_Ω, then there are only $\binom{m-p}{r+1-q}\binom{n-q}{r+1-q} + |\Omega|$ equations which generate $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ by Theorem 36. In particular, if one can find a fully known and full rank $r \times r$ submatrix of M_Ω, then there are only $(m-r)(n-r) + |\Omega|$ equations which generate $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$. This can significantly reduce the processing time needed to find the rank r completions of a partially known matrix M_Ω using algebraic methods.

Example 37. *Consider the partially known matrix given in Example 3,*

$$M(x,y,z,w) = \begin{bmatrix} 4 & 2 & 1 & x \\ 5 & 3 & y & 1 \\ 6 & z & 3 & 1 \\ w & 5 & 4 & 1 \end{bmatrix},$$

where x, y, z, and w are unknown. Note that the top-left submatrix $\left[\begin{smallmatrix} 4 & 2 \\ 5 & 3 \end{smallmatrix}\right]$ is nonsingular. So by Theorem 36, $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_2$ is equal to the zero set of all 3×3 minors containing the submatrix $\left[\begin{smallmatrix} 4 & 2 \\ 5 & 3 \end{smallmatrix}\right]$. That is, $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is *equal to the set of* $M(x,y,z,w)$ *such that*

$$\begin{vmatrix} 4 & 2 & 1 \\ 5 & 3 & y \\ 6 & z & 3 \end{vmatrix} = \begin{vmatrix} 4 & 2 & x \\ 5 & 3 & 1 \\ 6 & z & 1 \end{vmatrix} = \begin{vmatrix} 4 & 2 & 1 \\ 5 & 3 & y \\ w & 5 & 4 \end{vmatrix} = \begin{vmatrix} 4 & 2 & x \\ 5 & 3 & 1 \\ w & 5 & 1 \end{vmatrix} = 0.$$

These equations simplify to

$$\begin{aligned} -4yz + 12y + 5z - 12 &= 0 \\ 5xz - 18x - 4z + 14 &= 0 \\ 2yw - 20y - 3w + 33 &= 0 \\ -3xw + 25x + 2w - 18 &= 0. \end{aligned}$$

Solving this system of equations, we recover the two solutions from Example 3. There are sixteen 3×3 *submatrices of a* 4×4 *matrix up to permutation of rows and columns, which means there are sixteen equations which generate* $\overline{\mathcal{M}}_2$*. However, we were able to generate* $\overline{\mathcal{M}}_2 \cap \mathcal{A}_\Omega$ *using only four equations which generate* $\overline{\mathcal{M}}_2$*.*

1.10 Algebraic Combinatorics

Given a mask $\Omega \subset [m] \times [n]$, define the bipartite graph $G(\Omega) = (V, W, \Omega)$ as the graph with vertices $V = [m]$ and $W = [n]$, and edges Ω. Then the $m \times n$ adjacency matrix of $G(\Omega)$ is equal to $B(\Omega)$.

Example 38. *The mask*

$$B(\Omega_1) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

Corresponds to the bipartite graph Figure 1.9.

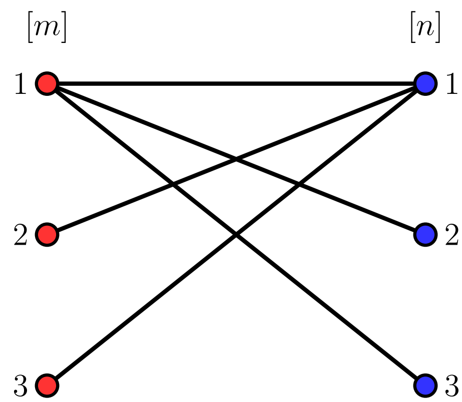


Figure 1.9: Bipartite graph $G(\Omega_1)$ from Example 38

Example 39. *The mask*

$$B(\Omega_2) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

Corresponds to the bipartite graph Figure 1.10.

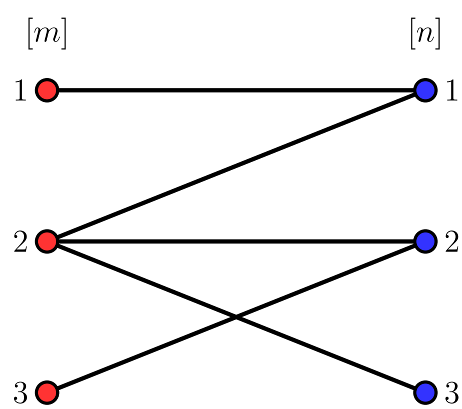


Figure 1.10: Bipartite graph $G(\Omega_2)$ from Example 39

Although a partially known matrix M_Ω may not be finitely completable in rank r, there may be individual positions have finitely many completions in a rank r completion of M_Ω. Let $P_{(i,j)}(X) = X_{ij}$ be the map which outputs the entry of X in position (i, j). Then in other words, the set $P_{(i,j)}(\Phi_\Omega^{-1}(M_\Omega))$ may be finite even if the set $\Phi_\Omega^{-1}(M_\Omega)$ is not finite.

Definition 40. *Given a partially known matrix M_Ω, a position $(i,j) \in [m] \times [n]$ is called* finitely completable in rank r *or* uniquely completable in rank r *if the set of values in position (i,j) is finite or unique respectively over all $X \in \Phi_\Omega^{-1}(M_\Omega)$.*

In other words, given M_Ω, the position (i,j) is finitely completable in rank r if

$$0 < \left|P_{(i,j)}(\Phi_\Omega^{-1}(M_\Omega))\right| < \infty$$

and is uniquely completable in r if

$$\left|P_{(i,j)}(\Phi_\Omega^{-1}(M_\Omega))\right| = 1.$$

Example 41. *Consider the partially know matrix $M_\Omega = \left[\begin{smallmatrix} 2 & 4 & \square \\ 1 & \square & \square \end{smallmatrix}\right]$. Then a matrix of the form $M = \left[\begin{smallmatrix} 2 & 4 & 2t \\ 1 & 2 & t \end{smallmatrix}\right]$ will be a rank one completion of M_Ω for any t, so M_Ω is not finitely completable in rank one. However, entry $(2,2)$ is uniquely completable in rank r. In other words, $P_{(2,2)}(X) = 2$ for all $X \in \Phi_\Omega^{-1}(M_\Omega)$.*

Given a mask Ω, different partially known matrices with the same mask may have different numbers of completions. However,

Definition 42. *Given a mask Ω, a position $(i,j) \in [m] \times [n]$ is called* finitely completable in rank r *or* uniquely completable in rank r *if, for a generic $M \in \overline{\mathcal{M}_r}$, the set of values in position (i,j) is finite or unique respectively over all $X \in \Phi_\Omega^{-1}(\Phi_\Omega(M))$.*

In other words, a position (i,j) is finitely completable in rank r if

$$0 < \left|P_{(i,j)}(\Phi_\Omega^{-1}(\Phi_\Omega(M)))\right| < \infty$$

and is uniquely completable in r if

$$\left|P_{(i,j)}(\Phi_\Omega^{-1}(\Phi_\Omega(M)))\right| = 1$$

for a generic $M \in \overline{\mathcal{M}_r}$.

We address the question of whether or not an unknown entry $(i, j) \in [m] \times [n]$ is finitely or uniquely completable. In other words, given an unknown entry $\square$ of M_Ω in position (i, j), are there only finitely many ways to complete $\square$ such that there exists a rank r completion of the resulting matrix? In general, for any continuous method of sampling, the question of whether or not the unknown entry $\square$ is uniquely completable depends only on the positions of the known entries in M_Ω with probability one (Király et al., 2015).

We introduce the set of all positions which are finitely completable, which is called the finitely completable closure.

Definition 43. *Given a mask* $\Omega \subset [m] \times [n]$, *the* rank r finitely completable closure $\mathrm{cl}_r(\Omega) \subset [m] \times [n]$ *is the set of positions which are finitely completable in* Ω.

To classify $\mathrm{cl}_r(\Omega)$, we introduce the following tools. Define the algebraic matrix multiplication map $\Upsilon : M_{m\times r} \times M_{n\times r} \to M_{m\times n}$ by $\Upsilon : (U, V) \mapsto UV^\top$. Note that every matrix in the image of this map has rank at most r. Moreover, $\overline{\mathcal{M}_r}$ is the image of the map Υ. Calculate the Jacobian J of Υ at point (U, V), the Jacobian has the following representation as an $mn \times r(m + n)$ matrix in terms of the Kronecker product, denoted $\otimes$, and identity matrices I_m and I_n.

$$J(U, V) = \begin{bmatrix} I_m \otimes v_1^\top & \\ \vdots & I_n \otimes U \\ I_m \otimes v_n^\top & \end{bmatrix}$$

where $v_j^\top$ is the jth row vector of V. Note that each row in $J(U, V)$ corresponds to some matrix entry (i, j). In particular, for a position $(i, j) \in [m] \times [n]$, the row $J_{(i,j)}(U, V)$ is defined to be the row of $J(U, V)$ corresponding to position (i, j). For a collection of known positions Ω, the matrix $J_\Omega(U, V)$ is defined to be the submatrix of $J(U, V)$ with rows corresponding to the known positions in Ω. We may use this definition to classify $\mathrm{cl}_r(\Omega)$.

Theorem 44. *(Király et al., 2015) Given a mask* Ω *and a rank* r*, we have*

$$\mathrm{cl}_r(\Omega) = \{(i,j) \in [m] \times [n] \mid J_{(i,j)}(U,V) \textit{ is in the row space of } J_\Omega(U,V)\}$$

for generic $UV^\top = M \in \overline{\mathcal{M}}_r$*.*

The linear independence of the rows of J_Ω is a generic property, and so it does not depend on the choice of generic M. Using Theorem 44, we may compute $\mathrm{cl}_r(\Omega)$ with Algorithm 6.

Algorithm 6: Finitely Completable Closure (Király et al., 2015)

Input: mask Ω, rank r
Result: rank r finite completable closure $\mathrm{cl}_r(\Omega)$
sample $U \in M_{m\times r}$ and $V \in M_{n\times r}$ from a continuous density;
calculate the Jacobian $J_\Omega(U,V)$;
for $(i,j) \in [m] \times [n]$ **do**
    if $J_{(i,j)}(U,V)$ *is in the row space of* $J_\Omega(U,V)$ **then**
        include (i,j) in $\mathrm{cl}_r(\Omega)$;
return $\mathrm{cl}_r(\Omega)$;

Sampling U and V from a continuous density will ensure that $M = UV^\top$ is generic with probability one. We now explicitly calculate $\mathrm{cl}_1(\Omega)$ in Example 41 with Algorithm 6.

Example 45. *Consider the case* $U = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ *and* $V^\top = \left[\begin{smallmatrix} 1 & 2 & 3 \end{smallmatrix}\right]$*. Then*

$$UV^\top = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix},$$

and

$$J(U,V) = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 2 \\ 0 & 3 & 0 & 0 & 1 \end{bmatrix}.$$

In this case we have,

$$\begin{aligned} J_{(1,1)} &= \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \end{bmatrix} \\ J_{(2,1)} &= \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ J_{(1,2)} &= \begin{bmatrix} 2 & 0 & 0 & 2 & 0 \end{bmatrix} \\ J_{(2,2)} &= \begin{bmatrix} 0 & 2 & 0 & 1 & 0 \end{bmatrix} \\ J_{(1,3)} &= \begin{bmatrix} 3 & 0 & 0 & 0 & 2 \end{bmatrix} \\ J_{(2,3)} &= \begin{bmatrix} 0 & 3 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Now let $\Omega = \{(1,1),(2,1),(1,2)\}$*. Then we have*

$$M_\Omega = \begin{bmatrix} 2 & 4 & \square \\ 1 & \square & \square \end{bmatrix}$$

and

$$J_\Omega = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 \end{bmatrix}.$$

Assuming U and V are generic, to check if entries $(2,2),(1,3)$, and $(2,3)$ are finitely completable, we need to check if $J_{(2,2)}$, $J_{(1,3)}$, and $J_{(2,3)}$ are in the row space of J_Ω. $J_{(2,2)}$ is in the row space of J_Ω because $\frac{1}{2}J_{(1,2)} + 2J_{(2,1)} - J_{(1,1)} = J_{(2,2)}$, so $(2,2) \in \mathrm{cl}_1(\Omega)$. However, $J_{(1,3)}$ and $J_{(2,3)}$ are not in the row space of J_Ω, so $\mathrm{cl}_1(\Omega) = \{(1,1),(1,2),(2,1),(2,2)\}$.

We are interested in the question of when finite completability implies unique completability. Like the finite completable closure, we introduce the set of positions which can be uniquely completed which is called the uniquely completable closure.

Definition 46. *Given a mask $\Omega \subset [m] \times [n]$, the* rank r uniquely completable closure $\mathrm{ucl}_r(\Omega) \subset [m] \times [n]$ *is the set of positions which are uniquely completable in Ω.*

To characterize the uniquely completable closure $\mathrm{ucl}_r(\Omega)$, we introduce the *stress* of a matrix.

Definition 47. *Given $U \in M_{m\times r}(\mathbb{C})$ and $V \in M_{n\times r}(\mathbb{C})$, a* rank r stress *of $M = UV^\top$ is a matrix that, as a vector, is in the kernel of the transpose of the Jacobian of U and V. In other words, a matrix $S \in M_{m\times n}(\mathbb{C})$ is a rank r stress of M if*

$$J(U,V)^\top \operatorname{vec}(S) = 0.$$

Given a mask Ω, the rank r stresses S such that $S_{ij} = 0$ for all $(i,j) \in \Omega^c$ are called the Ω-stresses of M in r. *The vector space of Ω-stresses of $M = UV^\top$ in r will be denoted $\Psi_M(\Omega)$.*

Given an Ω-stress S, note that if we vectorize S and remove zeros corresponding to unknown entries, then S is in the kernel of the submatrix $J_\Omega(U,V)^\top$. The maximum rank of the Ω-stresses is the property that allows us to test for unique completability.

Definition 48. *The* maximal Ω-stress rank of M in r *is defined as*

$$\rho_M(\Omega) = \max_{S\in\Psi_M(\Omega)} \operatorname{rank}(S)$$

If M is generic, the maximal stress rank $\rho_M(\Omega)$ depends only on Ω and r and not on M (Király et al., 2015). In this case, the generic Ω-stress rank in r is denoted $\rho(\Omega)$. We may now formulate the following theorem which gives conditions for when finite completability implies unique completability.

Theorem 49. *(Király et al., 2015) Given a mask Ω, if the generic Ω-stress rank in r satisfies the inequality*

$$\rho(\Omega) \geq \min(m, n) - r,$$

then $\mathrm{cl}_r(\Omega) = \mathrm{ucl}_r(\Omega)$.

Determining if finite completability implies unique completability relies on calculating the generic Ω-stress $\rho(\Omega)$, which can be done by using Algorithm 7.

Algorithm 7: Generic Stress Rank (Király et al., 2015)

Input: mask Ω, rank r
Result: generic stress rank $\rho(\Omega)$ of Ω in rank r

1. sample $U \in M_{m\times r}$ and $V \in M_{n\times r}$ from a continuous density;
2. calculate Jacobian $J_\Omega(U, V)$;
3. sample a random vector $\mathrm{vec}(S) \in \mathbb{R}^{|\Omega|}$ in the kernel of $J_\Omega(U, V)^\top$ from a continuous density;
4. reshape $\mathrm{vec}(S)$ into an $m \times n$ matrix S with entries in Ω corresponding to entries in $\mathrm{vec}(S)$ and entries in Ω^c as zeros;
5. output $\rho(\Omega) = \mathrm{rank}(S)$;

Again, sampling $\mathrm{vec}(S)$ from a continuous density will ensure $\mathrm{vec}(S)$ is generic with probability one. Once $\rho(\Omega)$ is calculated, we simply check if $\rho(\Omega) \geq \min(m, n) - r$ to see if $\mathrm{cl}_r(\Omega) = \mathrm{ucl}_r(\Omega)$ by Theorem 49.

1.11 Algebraic Topology

In this section we discuss some topology properties of the spaces $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$. We start with some definitions from algebraic topology.

Definition 50. *An n-dimensional* simplex *is an n-dimensional polytope which is the convex hull of $n + 1$ points. A* simplicial complex *is a topological space V which is divided into simplices such that every face of* *a simplex in V is also a simplex in V and the non-empty intersection of any two simplices is a face of each simplex.*

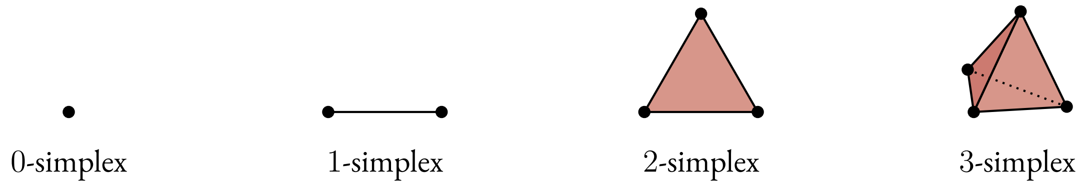


Figure 1.11: Examples of n-simplices for $n = 0, 1, 2, 3$

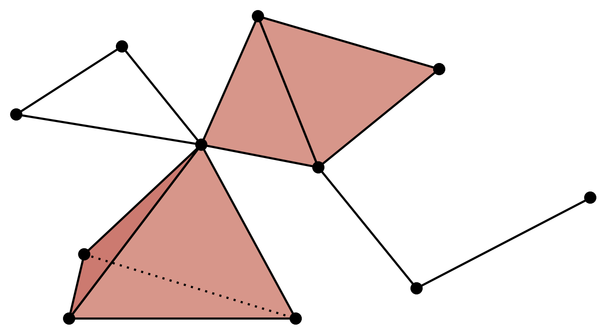


Figure 1.12: Example of a simplicial complex

Definition 51. *Given a simplicial complex V, the* homology groups $H_k(V)$, $k = 0, 1, \ldots$ *are abelian groups which may be used to count the number of holes in V. The kth* Betti number *of V is equal to the number of k-dimensional holes in V, and is denoted $h_k(V)$. More formally, assuming V has finitely many k-dimensional holes, then $H_k(V) \otimes \mathbb{R} \cong \mathbb{R}^d$ for some d, in which case $h_k(V) = d$. Here $\otimes$ denotes the tensor product of abelian groups as modules over $\mathbb{Z}$.*

Example 52. *$h_0(V)$ is equal to the number of connected components of V, so if V is a collection of N points, then $h_0(V) = N$. $h_1(V)$ is equal to the number of one dimensional holes in V. So if V is a collection of N loops, then $h_1(V) = N$.*

Definition 53. *A sequence of homomorphisms*

$$\cdots \xrightarrow{f_{i-1}} G_i \xrightarrow{f_i} G_{i+1} \xrightarrow{f_{i+1}} G_{i+2} \xrightarrow{f_{i+2}} \cdots$$

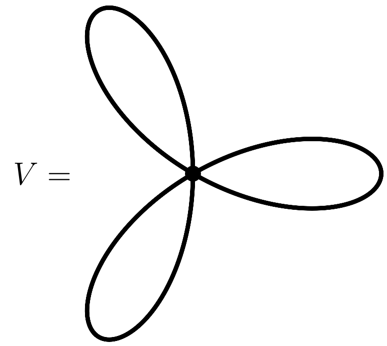


Figure 1.13: A collection of three loops glued at a point has three one-dimensional holes, so $h_1(V) = 3$. The figure is connected, so $h_0(V) = 1$.

of abelian groups G_i is called exact *if the kernel of f_i is equal to the image of f_{i-1} for all i.*

We now introduce the Mayer-Vietoris sequence which relates the homology groups of simplicial complexes.

Theorem 54. *(Hatcher, 2002) Let V and W be simplicial complexes such that $V \cup W$ and $V \cap W$ are also simplicial complexes. There exists an exact sequence of the form*

$$\cdots \longrightarrow H_1(V) \oplus H_1(W) \longrightarrow H_1(V \cup W) \longrightarrow H_0(V \cap W) \longrightarrow H_0(V) \oplus H_0(W) \longrightarrow H_0(V \cup W) \longrightarrow 0$$

which is called the Mayer-Vietoris sequence.

Here $\oplus$ denotes the direct sum of abelian groups. We use the Mayer-Vietoris sequence to obtain a relationship between the Betti numbers $h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r)$ and $h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$. First, we introduce the following theorem.

Theorem 55. *((Hironaka, 1975) p.180) Affine algebraic varieties V and W may be triangulated so that V, W, $V \cup W$ and $V \cap W$ are simplicial complexes.*

This means that we may apply the Mayer-Vietoris sequence to $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$.

Theorem 56. *Given a partially known matrix M_Ω with at least one rank at most r completion,*

$$h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) + 1.$$

Proof. Because M_Ω has at least one rank at most r completion, then $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is nonempty. By Theorem 54 and Theorem 55, we have an exact sequence of the form

$$\cdots \longrightarrow H_1(\mathcal{A}_\Omega) \oplus H_1(\overline{\mathcal{M}}_r) \longrightarrow H_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) \longrightarrow H_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) \longrightarrow H_0(\mathcal{A}_\Omega) \oplus H_0(\overline{\mathcal{M}}_r) \longrightarrow H_0(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) \longrightarrow 0 \tag{1.4}$$

Because scaling does not change the rank of a matrix, $\overline{\mathcal{M}}_r$ deformation retracts to the singleton $\{0\}$ via the map

$$F : \overline{\mathcal{M}}_r \times [0,1] \to \overline{\mathcal{M}}_r$$
$$F(X, \alpha) = (1-\alpha)X$$

so $\overline{\mathcal{M}}_r$ is homotopy equivalent to a point. Similarly, $\mathcal{A}_\Omega$ is a linear variety so it deformation retracts to the singleton $\{M_\Omega\}$ via the map

$$G : \mathcal{A}_\Omega \times [0,1] \to \mathcal{A}_\Omega$$
$$G(X, \alpha) = (1-\alpha)X + \alpha M_\Omega.$$

Because homology groups are invariant under homotopy equivalence, both $\overline{\mathcal{M}}_r$ and $\mathcal{A}_\Omega$ have the homology of a point. This implies that

$$H_1(\overline{\mathcal{M}}_r) \cong 0 \qquad H_1(\mathcal{A}_\Omega) \cong 0$$
$$H_0(\overline{\mathcal{M}}_r) \cong \mathbb{Z} \qquad H_0(\mathcal{A}_\Omega) \cong \mathbb{Z}.$$

Because $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$ are both connected and have non-empty intersection, their union is also connected, so

$$H_0(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) \cong \mathbb{Z}.$$

Therefore, Sequence (1.4) simplifies to

$$0 \longrightarrow H_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) \longrightarrow H_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) \longrightarrow \mathbb{Z}^2 \longrightarrow \mathbb{Z} \longrightarrow 0. \tag{1.5}$$

Because $\mathbb{R}$ is a flat $\mathbb{Z}$ module, we may take the tensor product of Sequence (1.5) with $\mathbb{R}$ to obtain

$$0 \longrightarrow \mathbb{R}^{h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)} \longrightarrow \mathbb{R}^{h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r)} \longrightarrow \mathbb{R}^2 \longrightarrow \mathbb{R} \longrightarrow 0. \tag{1.6}$$

By the general rank-nullity theorem, if

$$0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow \cdots \longrightarrow V_k \longrightarrow 0$$

is an exact sequence of vector spaces V_i, then $\sum_{i=1}^{k}(-1)^i \dim(V_i) = 0$. Applying this theorem to Sequence (1.6) we get $h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) + 1$. $\square$

$h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r)$ is equal to the number of connected components of $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$, so if M_Ω is finitely completable in rank r, then $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is a finite collection of points. So

$$h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = \left|\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r\right| = N,$$

where N is the number of distinct rank r completions of M_Ω. So $N = h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) + 1$.

Example 57. *Consider Example 3 where there are two intersection points M_1 and M_2 in $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$. Because there are two connected components, $h_0(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = 2$, which implies $h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) = 1$ by Theorem 56. This means that there is a single one-dimensional hole in $\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r$. See Figure 1.14 for visual intuition.*

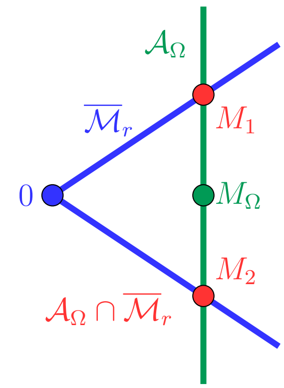


Figure 1.14: A schematic of $\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r$. There are two points in $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ and one hole in $\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r$.

If M_Ω is finitely completable in rank r and it is possible to calculate $h_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$, then we may find the number of rank r completions of M_Ω with Theorem 56. One strategy may be to calculate the fundamental group of $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ because its abelianization is equal to $H_1(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$.

Definition 58. *The* Euler characteristic *of a simplicial complex V, is equal to the alternating sum of the Betti numbers of V, and is denoted $\chi(V)$. That is,*

$$\chi(V) = \sum_{k=0}^{\infty} (-1)^k h_k(V).$$

Because the kth homology group of V is trivial for all $k > \dim(V)$ (Hatcher, 2002), the kth Betti number is equal to zero for all $k > \dim(V)$, so the infinite sum in Definition 58 is actually a finite sum. Moreover because Betti numbers are homotopy invariant, the Euler characteristic is also homotopy invariant.

Example 59. *In fig. 1.13, $h_0(V) = 1$, $h_1(V) = 3$, and $h_k(V) = 0$ for all $k \geq 2$, so V has Euler characteristic $\chi(V) = -2$.*

Like Theorem 54, given simplicial complexes V and W, there is a relationship between the Euler characteristics of V, W, $V \cap W$, and $V \cup W$.

Theorem 60. *(Hatcher, 2002) Given simplicial complexes V and W such that $V \cup W$ and $V \cap W$ are also simplicial complexes, the Euler characteristic satisfies the inclusion-exclusion principle*

$$\chi(V \cup W) = \chi(V) + \chi(W) - \chi(V \cap W).$$

Like Theorem 56, we have the following theorem relating $\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r)$ and $\chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$.

Theorem 61. *Given a partially known matrix M_Ω with at least one rank at most r completion, we have*

$$\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = 2 - \chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r).$$

Proof. Because $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$ are homotopy equivalent to a point, $\chi(\overline{\mathcal{M}}_r) = \chi(\mathcal{A}_\Omega) = 1$. Applying Theorem 60 to $\mathcal{A}_\Omega$ and $\overline{\mathcal{M}}_r$, we have

$$\begin{aligned}\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) &= \chi(\mathcal{A}_\Omega) + \chi(\overline{\mathcal{M}}_r) - \chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r) \\ &= 2 - \chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r).\end{aligned}$$

□

Again, if M_Ω is finitely completable in rank r, let $N = \left|\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r\right|$ be the number of rank r completions of M_Ω. Then because $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is a finite collection of N points, $\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = N$, so $N = 2 - \chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$. If M_Ω is finitely completable in rank r and it is possible to calculate $\chi(\mathcal{A}_\Omega \cup \overline{\mathcal{M}}_r)$, then we may find the number of rank r completions of M_Ω with Theorem 61.

Example 62. *Consider M_Ω as given in Example 57. Because $\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r$ is a collection of two points, we have $\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = 2$, which implies $\chi(\mathcal{A}_\Omega \cap \overline{\mathcal{M}}_r) = 0$ by Theorem 61.*

Chapter 2

Maximum Volume Principle

In several matrix analysis problems, knowledge of a quality submatrix A of a large matrix X is required. In general, A should not be close to singular, so one way to measure the quality of A is by modulus of the determinant, which is called the *volume*.

Definition 63. *The absolute value of the determinant of a matrix X, is referred to as the* volume *of X and is denoted* $\text{vol}(X) = |\det(X)|$.

The volume of a linear map X is the factor by which the measure of a measurable set changes under X ((Aliprantis & Burkinshaw, 1998) p.389). Moreover, it is equal to the product of the singular values of X.

2.1 Schur Complement

We introduce the Schur complement as an example of when knowledge of a quality submatrix is desired.

Definition 64. *Let $X \in M_{m\times n}$ be an $m \times n$ matrix. By permuting rows and columns, suppose X has the block structure $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ where A is $r \times r$ nonsingular, B is $r \times (n-r)$, C is $(m-r) \times r$, and D is $(m-r) \times (n-r)$. The* Schur complement *of X with respect to A is denoted S_A and is defined as*

$$S_A = D - CA^{-1}B.$$

In general, A may be any $r \times r$ nonsingular submatrix, in which case we permute the rows and columns of X so that A is in the top-left corner. If X is square, the Schur complement of X with respect to A has a useful formula relating $\det(X)$ and $\det(A)$ called the *Schur determinant identity*.

Lemma 65. *(Zhang, 2006) Given an $n \times n$ matrix X, for any $r \times r$ nonsingular submatrix A in X, we have*

$$\det(X) = \det(A)\det(S_A).$$

Moreover, by taking the absolute value, we also have

$$\operatorname{vol}(X) = \operatorname{vol}(A)\operatorname{vol}(S_A).$$

Proof. (Zhang, 2006) Let $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ be an $n \times n$ block matrix where A is $r \times r$ and nonsingular. Then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & I \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

where I are identity matrices of the appropriate size. Taking the determinant we get the desired result. □

The rank of the Schur complement has a useful relation the the rank of X.

Lemma 66. *(Zhang, 2006) Given an $m \times n$ matrix X, for any nonsingular $r \times r$ submatrix A in X, $S_A = 0$ if and only if $\operatorname{rank}(X) = r$. More generally,*

$$\operatorname{rank}(X) = \operatorname{rank}(S_A) + r.$$

2.2 Skeleton Decomposition

Let $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ be an $m \times n$ block matrix where A is an $r \times r$ nonsingular submatrix. By Lemma 66, if $\operatorname{rank}(X) = r$, then $D = CA^{-1}B$, and so $X = \left[\begin{smallmatrix} A & B \\ C & CA^{-1}B \end{smallmatrix}\right]$. This is directly related to the matrix completion Example 32.

We may use the skeleton decomposition to decompose X as a product of an $m \times r$ matrix, an $r \times r$ matrix, and an $r \times n$ matrix.

Definition 67. *Let X be an $m \times n$ rank r matrix with an $r \times r$ nonsingular submatrix A. After permutation of rows and columns, suppose X has the block structure $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ for corresponding B, C, and D. Then*

$$X = \begin{bmatrix} A \\ C \end{bmatrix} A^{-1} \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} A & B \\ C & CA^{-1}B \end{bmatrix}$$

is the rank r skeleton decomposition *of X with respect to A.*

What if X is not exactly rank r, but only approximately rank r? Then we may obtain a rank r approximation X_r of X using the skeleton decomposition.

Definition 68. *Let X be an $m \times n$ matrix with an $r \times r$ nonsingular submatrix A. After permutation of rows and columns, suppose X is in the block form $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. Then*

$$X_r = \begin{bmatrix} A \\ C \end{bmatrix} A^{-1} \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} A & B \\ C & CA^{-1}B \end{bmatrix}$$

is the rank r skeleton decomposition *of X with respect to A. The skeleton decomposition is also sometimes called the CUR* decomposition*.*

Note that $\operatorname{rank}(X_r) = r$. Moreover, the row space of X_r is spanned by the first r rows, and the columns space is spanned by the first r columns. Also note that X_r may be stored by only storing A, B, and C, which amounts to $(m+n)r - r^2$ entries. On the other hand, X has mn entries total, so the *compression ratio* of the skeleton decomposition is equal to

$$\frac{(m+n)r - r^2}{mn}.$$

The skeletal decomposition and the Schur complement are directly related. Let X_r be a rank r skeleton decomposition of X with respect to a submatrix A. Then

$$\begin{aligned}\|X - X_r\|_{\max} &= \left\| \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} A & B \\ C & CA^{-1}B \end{bmatrix} \right\|_{\max} \\ &= \left\| \begin{bmatrix} 0 & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \right\|_{\max} \\ &= \left\| D - CA^{-1}B \right\|_{\max} \\ &= \|S_A\|_{\max}\end{aligned}$$

where $\|X\|_{\max} = \max_{i,j} |X_{ij}|$ is the max norm. So the error of the skeleton decomposition of X with respect to A in the max norm is equal to the max norm of the Schur complement of X with respect to A. Note that the error $\|X - X_r\|_{\max} = 0$ if and only if $\operatorname{rank}(X) = r$ by Lemma 66.

The error of the rank r skeleton decomposition to the original matrix is in general larger than the best rank r decomposition obtained with the singular value decomposition by Theorem 11. However, the skeleton decomposition has the benefit being a rational function of the actual entries of M, making the skeleton decomposition much faster to compute.

The error of the skeleton decomposition will depend on the choice of the $r \times r$ submatrix A. In particular, the quality of the skeleton decomposition with respect to A depends on the volume of A. It is desirable for the volume of A to be as large as possible.

Theorem 69. *(Goreinov et al., 2010) The error of the skeleton decomposition with respect to the max norm is minimized over all choices of* $r \times r$ *submatrices* A *when* $\operatorname{vol}(A)$ *is maximized. Let* X_r *be the rank* r *skeleton decomposition of* X *with respect to a submatrix* A*. Then when* A *is chosen with maximum volume over all choices of* $r \times r$ *submatrices, we have the inequality*

$$\|X - X_r\|_{\max} \leq (r+1)\sigma_{r+1}(X)$$

where $\sigma_{r+1}(X)$ is the $(r+1)$st singular value of X.

Recall that $\sigma_{r+1}(X)$ is the error of X to a best rank r approximation of X in the spectral norm.

We denote an $r \times r$ submatrix of maximum volume over all $r \times r$ submatrices of X by $X_{\max}$. Then

$$\operatorname{vol}(X_{\max}) = \max_{|I|=|J|=r} \operatorname{vol}(X_{I,J}).$$

Although it is desirable to find an $r \times r$ submatrix of maximum volume, it is not always feasible because finding $X_{\max}$ is an NP-hard problem (Civril & Magdon-Ismail, 2009). It is much easier to find near optimal volume submatrices by searching for submatrices which have locally maximum volume as opposed to globally maximum volume.

To make this concept more precise, we first introduce the following lemma.

Lemma 70. *Let X be an $m \times r$ matrix with $r \leq m$. Let A be an $r \times r$ nonsingular submatrix of X, and let $Y = XA^{-1}$. Let B equal A with the jth row of A replaced with the ith row of X. Then*

$$\operatorname{vol}(B) = |Y_{ij}| \operatorname{vol}(A)$$

Proof. Note that $\frac{\operatorname{vol}(B)}{\operatorname{vol}(A)} = \operatorname{vol}(BA^{-1})$. Let $Z = BA^{-1}$, we want to show $\operatorname{vol}(Z) = |Y_{ij}|$. We have

$$Z_{pq} = \begin{cases} Y_{iq} & \text{if } p = j \\ \delta_{pq} & \text{if } p \neq j \end{cases}$$

where Z_{pq} is the entry of Z in position (p, q), and δ_{pq} is the Kronecker delta which is equal to one if $p = q$ and equal to zero otherwise. By swapping first and jth row, and the first and jth columns of Z, we get an upper diagonal matrix with entry in position $(1, 1)$ equal to Y_{ij} and with the rest of the diagonal entries equal to 1. Therefore, $\operatorname{vol}(BA^{-1}) = |Y_{ij}|$. □

Lemma 70 motivates the following definition.

Definition 71. *An $r \times r$ nonsingular submatrix $\hat{X}$ of an $m \times r$ matrix X, is called* dominant *if all entries of $X\hat{X}^{-1}$ are at most one in modulus.*

Alternatively, because $X\hat{X}^{-1}$ contains the identity matrix, a submatrix $\hat{X}$ is dominant if

$$\left\| X\hat{X}^{-1} \right\|_{\max} = 1.$$

It follows from Lemma 70, that we may not increase the volume of a dominant submatrix $\hat{X}$ by swapping a row in $\hat{X}$ with a row in X. It is in this sense that dominant submatrices $\hat{X}$ have locally maximal volume. Analogous to the statement that global maximums are also local maximums, we have the following theorem.

Theorem 72. *(Goreinov et al., 2010) Given an $m \times r$ matrix X with $r \leq m$, an $r \times r$ submatrix of maximum volume $X_{\max}$ is a dominant submatrix.*

Proof. Suppose $X_{\max}$ is not dominant. Then there exists an entry Y_{ij} of $Y = XX_{\max}^{-1}$ such that $|Y_{ij}| > 1$. Let B be equal to $X_{\max}$ with the jth row of $X_{\max}$ replaced with the ith row of X. Then by Lemma 70,

$$\mathrm{vol}(X_{\max}) < |Y_{ij}| \, \mathrm{vol}(X_{\max}) = \mathrm{vol}(B).$$

This contradicts the assumption that $X_{\max}$ is a maximum volume submatrix, so $X_{\max}$ is dominant. $\square$

Dominant submatrices are useful because they are easier to search for than maximum volume submatrices. Moreover, they and are not too far off in volume from a maximum volume submatrix. To show this, first recall Hadamard's inequality.

Lemma 73. *(Hadamard, 1893) Given an $n \times n$ matrix $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ where x_i is the ith column of X, Hadamard's inequality states that*

$$\mathrm{vol}(X) \leq \prod_{i=1}^{m} \|x_i\| \, .$$

Moreover, if $|X_{ij}| \leq N$ *for all* $i \leq n$ *and* $j \leq n$*, then* $\text{vol}(X) \leq N^n n^{n/2}$.

Now we may state the following theorem relating the volume of dominant submatrices and maximum volume submatrices.

Theorem 74. *(Goreinov et al., 2010) Given an* $m \times r$ *matrix* X*, let* $X_{\max}$ *be an* $r \times r$ *maximum volume submatrix, and let* $\hat{X}$ *be an* $r \times r$ *dominant submatrix. Then we have the inequality*

$$\text{vol}(X_{\max}) \leq r^{r/2}\,\text{vol}(\hat{X})$$

Proof. (Goreinov et al., 2010) Let $Z = X_{\max}\hat{X}^{-1}$. Then because $\hat{X}$ is dominant, $|Z_{ij}| \leq 1$ for all $i \leq r$ and $j \leq r$. So by Lemma 73,

$$\frac{\text{vol}(X_{\max})}{\text{vol}(\hat{X})} = \text{vol}(Z) \leq r^{r/2}.$$

□

It is useful to generalize the definition of an $r \times r$ dominant submatrix of matrices of size $m \times r$ to matrices of any size $m \times n$.

Definition 75. *Given an* $m \times n$ *matrix* X*, an* $r \times r$ *nonsingular submatrix* $\hat{X}$ *is called* dominant *if it is dominant in its respective rows and columns in the sense of Definition 71.*

2.3 Maximum Volume Algorithm

One nice feature of dominant submatrices is that they are easy to find. In practice, given an $m \times r$ matrix X, we only need to find a close to dominant $r \times r$ submatrix $\hat{X}$ such that

$$\left\| X\hat{X}^{-1} \right\|_{\max} \leq 1 + \epsilon \tag{2.1}$$

for some tolerance ϵ. This may significantly reduce the number of iterations needed to compute $\hat{X}$, while practically keeping the volume the same because if Equation (2.1) holds, replacing a row of $\hat{X}$ with a row of X will only increase the volume by at most a factor of $1+\epsilon$.

We start by introducing the standard algorithm for finding an $r \times r$ dominant submatrix of an $m \times r$ matrix which is called the *maximum volume* or *maxvol* algorithm.

Algorithm 8: Maximum Volume (Goreinov et al., 2010)

```
Input: m × r matrix X, r × r nonsingular submatrix A_0, tolerance ε ≥ 0
Result: A_k a close to dominant submatrix of X.
for k = 0, 1, ... do
    Y_k = X A_k^{-1};
    find a largest in modulus entry of Y_k called (Y_k)_ij with index (i, j);
    if |(Y_k)_ij| ≤ 1 + ε then
        return A_k;
    else
        A_{k+1} equals A_k with the jth row replaced with the ith row of X;
```

The the volume of the submatrices A_k produced by Algorithm 8 increases until a close to dominant submatrix is reached. In other words, we have the following theorem.

Theorem 76. *(Goreinov et al., 2010) The sequence of matrices $\{A_k\}$ given by Algorithm 8 has increasing volume which is bounded from above, so Algorithm 8 converges. That is,*

$$\operatorname{vol}(A_k) < \operatorname{vol}(A_{k+1}).$$

Proof. By Lemma 70, on the kth step,

$$\operatorname{vol}(A_k) < |(Y_k)_{ij}| \operatorname{vol}(A_k) = \operatorname{vol}(A_{k+1})$$

because $|(Y_k)_{ij}| > 1$. Moreover, there are only finitely many submatrices in X, so the volume of the submatrices in X are bounded. For a more explicit upper bound, we may use Lemma 73. □

Note that A_{k+1} has the largest volume obtainable from A_k by swapping one row in A_k with a row in X. This is because $(Y_k)_{ij}$ is chosen so that it is the largest entry in modulus over all indices (i, j).

We can generalize Algorithm 8 to find an $r \times r$ dominant submatrix of an $m \times n$ matrix X. We may search for the largest in modulus entry of both $Y_k = X_{[m],J} A_k^{-1}$, and $Z_k = A_k^{-1} X_{I,[n]}$ at each iteration, where A_k is the submatrix with entries in positions $I \times J$ in X.

Algorithm 9: Two Directional Maximum Volume

Input: $m \times n$ matrix X, $r \times r$ nonsingular submatrix A_0, tolerance $\epsilon \geq 0$
Result: A_k a close to dominant submatrix in X.
for $k = 0, 1, \ldots$ **do**
  let I and J be the row and column indices of $A_k = X_{I,J}$;
  $Y_k = X_{[m],J} A_k^{-1}$;
  find a largest in modulus entry of Y_k called $(Y_k)_{ij}$ in position (i, j);
  $Z_k = A_k^{-1} X_{I,[n]}$;
  find a largest in modulus entry of Z_k called $(Z_k)_{pq}$ in position (p, q);
  if $\max(|(Y_k)_{ij}|, |(Z_k)_{pq}|) \leq 1 + \epsilon$ **then**
    return A_k;
  else
    if $|(Y_k)_{ij}| \geq |(Z_k)_{pq}|$ **then**
      A_{k+1} equals A_k with the jth row of A_k replaced with the ith row of $X_{[m],J}$;
    else
      A_{k+1} equals A_k with the pth column of A_k replaced with the qth column of $X_{I,[n]}$;

However, Algorithm 9 calculates both $Y_k = X_{[m],J} A_k^{-1}$ and $Z_k = A_k^{-1} X_{I,[n]}$ at each iteration. To speed up the processing time, we may consider an alternating maximum volume algorithm where we alternate between optimizing rows and columns to make one such calculation at each iteration.

Because the sequence $\{\text{vol}(A_k)\}$ is again increasing and bounded above, both Algorithm 9 and Algorithm 10 converge to a dominant submatrix.

Note that because the the alternating maximum volume algorithm only searches either rows or columns at each iteration, it may not always find the largest volume submatrix obtainable with one row permutation or one column permutation. For example, consider the matrix $X = \left[\begin{smallmatrix} 2 & 3 \\ 1 & 4 \end{smallmatrix}\right]$. Let $r = 1$, and let A_0 be the 1×1 submatrix of X consisting of the entry 1 in position $(2, 1)$. Then Algorithm 9 converges to the

Algorithm 10: Alternating Maximum Volume

Input: $m \times n$ matrix X, $r \times r$ nonsingular submatrix A_0, tolerance $\epsilon \geq 0$
Result: A close to dominant submatrix in X.
initialize $(Z_{-1})_{pq} = \infty$;
for *k=0,1,…* **do**
  let I and J be the row and column indices of $A_k = X_{I,J}$;
  $Y_k = X_{[m],J} A_k^{-1}$;
  find a largest in modulus entry of Y_k called $(Y_k)_{ij}$ in position (i, j);
  if $|(Y_k)_{ij}| > 1 + \epsilon$ **then**
    B_k equals A_k with the jth row of A_k replaced with the ith row of $X_{[m],J}$;
  else
    if $|(Z_{k-1})_{pq}| \leq 1 + \epsilon$ **then**
      return A_k;
    else
      $B_k = A_k$;
  $Z_k = B_k^{-1} X_{I,[n]}$;
  find a largest in modulus entry of Z_k called $(Z_k)_{pq}$ in position (p, q);
  if $|(Z_k)_{pq}| > 1 + \epsilon$ **then**
    A_{k+1} equals B_k with the pth column of B_k replaced with the qth column of $X_{I,[n]}$;
  else
    if $|(Y_k)_{ij}| \leq 1 + \epsilon$ **then**
      return B_k;
    else
      $A_{k+1} = B_k$;

1×1 submatrix consisting of the entry 4 in position $(2, 2)$ in one step, and Algorithm 10 converges in three iterations with the same initialization. More generally, we have the following example.

Example 77. *Let*

$$X = \begin{bmatrix} 2 & 3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 4 & 5 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 6 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2n-4 & 2n-3 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 2n-2 & 2n-1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 2n \end{bmatrix}$$

be the $n \times n$ matrix such that the kth diagonal entry is equal to $2k$, the kth upper diagonal entry is equal to $2k+1$, and the entry in position $(n, 1)$ is equal to one. Let A_0 be the 1×1 submatrix consisting of the entry 1 in position $(n, 1)$. Then Algorithm 9 converges to the 1×1 submatrix consisting of the entry $2n$ in position (n, n) in one step, and Algorithm 10 converges in $2n - 1$ iterations with the same initialization.

2.4 Greedy Maximum Volume Algorithm

Note that Algorithm 8, Algorithm 9, and Algorithm 10, only permute one row or column at each iteration. We may reduce the processing time needed to find a dominant submatrix by potentially swapping more rows at each iteration. We call such an algorithm a *greedy maximum volume algorithm*. Note that the maximum volume algorithm, Algorithm 8, is already a type of greedy algorithm in the sense that local optimizations are made at each iteration.

The greedy maximum volume algorithm for finding an $r \times r$ dominant submatrix of an $m \times r$ matrix is similar to Algorithm 8. The difference is instead of swapping one row each iteration, we may swap two or more rows. The determinant of a small submatrix is also calculated at each iteration to ensure that the volume is increasing.

First, we describe the simplified case where we swap at most two rows per step. We call this algorithm the simple greedy maximum volume algorithm.

Algorithm 11: Simple Greedy Maximum Volume

Input: $m \times r$ matrix X, $r \times r$ nonsingular submatrix A_0, tolerance $\epsilon \geq 0$
Result: A_k a close to dominant submatrix of X.
for $k = 0, 1, \ldots$ **do**
    $Y_k = XA_k^{-1}$;
    find a largest in modulus entry of Y_k called $(Y_k)_{ij}$ in position (i, j);
    if $|(Y_k)_{ij}| \leq 1 + \epsilon$ **then**
        return A_k;
    else
        find a largest in modulus entry of Y_k not in row i or column j called $(Y_k)_{pq}$ in position (p, q);
        $V_k = \begin{bmatrix} (Y_k)_{ij} & (Y_k)_{iq} \\ (Y_k)_{pj} & (Y_k)_{pq} \end{bmatrix}$;
        if $\text{vol}(V_k) > |(Y_k)_{ij}|$ **then**
            A_{k+1} equals A_k with the jth row replaced with the ith row of X, and the qth row replaced with the pth row of X;
        else
            A_{k+1} equals A_k with the jth row replaced with the ith row of X;

Like Algorithm 8, the volume of the submatrices A_k produced by Algorithm 11 increases until a close to dominant submatrix is reached, so it converges. To prove that the sequence of submatrices $\{A_k\}$ has increasing volume, we first generalize Lemma 70.

Lemma 78. *Let X be an $m \times r$ matrix with $r \leq m$. Let A be an $r \times r$ nonsingular submatrix of X, and let $Y = XA^{-1}$. Given a set of s pairs of indices $\{(i_1, j_1), \ldots, (i_s, j_s)\} \subset [m] \times [r]$ where $s \leq r$, let B equal A with row j_α row of A replaced with row i_α of X for all $1 \leq \alpha \leq s$. Define the matrix*

$$V = \begin{bmatrix} Y_{i_1 j_1} & Y_{i_1 j_2} & \ldots & Y_{i_1 j_s} \\ Y_{i_2 j_1} & Y_{i_2 j_2} & \ldots & Y_{i_2 j_s} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{i_s j_1} & Y_{i_s j_2} & \ldots & Y_{i_s j_s} \end{bmatrix}.$$

Then

$$\mathrm{vol}(B) = \mathrm{vol}(V)\,\mathrm{vol}(A)$$

Proof. Note that $\frac{\mathrm{vol}(B)}{\mathrm{vol}(A)} = \mathrm{vol}(BA^{-1})$. Let $Z = BA^{-1}$, we want to show $\mathrm{vol}(Z) = \mathrm{vol}(V)$. We have

$$Z_{pq} = \begin{cases} Y_{i_\alpha q} & \text{if } p = j_\alpha \text{ for some } \alpha \\ \delta_{pq} & \text{otherwise} \end{cases}$$

where Z_{pq} is the entry of Z in position (p, q), and δ_{pq} is the Kronecker delta which is equal to one if $p = q$ and equal to zero otherwise. Swap rows α and j_α, and columns α and j_α of Z for all α. Then Z is in the block upper diagonal form

$$Z = \begin{bmatrix} V & W \\ 0 & I \end{bmatrix}$$

for some submatrix W, where I is the $(r-s)\times(r-s)$ identity matrix. So we have $\mathrm{vol}(Z) = \mathrm{vol}(V)$. □

We now show that Algorithm 11 converges.

Theorem 79. *The sequence of matrices $\{A_k\}$ produced by Algorithm 11 has increasing volume which is bounded above, so Algorithm 11 converges. That is,*

$$\mathrm{vol}(A_k) < \mathrm{vol}(A_{k+1}).$$

Proof. First, suppose only one row in A_k is swapped on step k. Then by Lemma 70,

$$\mathrm{vol}(A_k) < |(Y_k)_{ij}|\,\mathrm{vol}(A_k) = \mathrm{vol}(A_{k+1})$$

because $|(Y_k)_{ij}| > 1$.

Now suppose two rows in A_k are swapped on step k. Then by Lemma 78 when $s = 2$,

$$\text{vol}(A_k) < \text{vol}(V_k)\,\text{vol}(A_k) = \text{vol}(A_{k+1})$$

because $\text{vol}(V_k) > |(Y_k)_{ij}| > 1$. Moreover, there are only finitely many submatrices in X, so the volume of the submatrices in X are bounded. $\square$

Note that when two rows are swapped in Algorithm 11, A_{k+1} has the largest volume obtainable from A_k when $\text{vol}(V_k)$ is maximized.

$$\text{vol}(V_k) = |(Y_k)_{ij}(Y_k)_{pq} - (Y_k)_{iq}(Y_k)_{pj}|$$

may be small even if $(Y_k)_{ij}$ and $(Y_k)_{pq}$ are large. In general,

We now define the greedy maximum volume algorithm, Algorithm 12. The greedy maximum volume algorithm is similar to Algorithm 11, but we may swap at most r rows per step as opposed to at most 2 rows.

Theorem 80. *The sequence of matrices $\{A_k\}$ given by Algorithm 12 has increasing volume which is bounded above, so Algorithm 12 converges. That is,*

$$\text{vol}(A_k) < \text{vol}(A_{k+1}).$$

Proof. Suppose s rows in A_k are swapped on step k. Then by Lemma 78,

$$\text{vol}(A_k) < \text{vol}(V_{k,s})\,\text{vol}(A_k) = \text{vol}(A_{k+1})$$

because

$$1 < \text{vol}(V_{k,1}) < \cdots < \text{vol}(V_{k,s-1}) < \text{vol}(V_{k,s}).$$

Algorithm 12: Greedy Maximum Volume Algorithm

Input: $m \times r$ matrix X, $r \times r$ nonsingular submatrix A_0, tolerance $\epsilon \geq 0$

Result: A_k a close to dominant submatrix of X

for $k = 0, 1, \ldots$ **do**

  $Y_k = XA_k^{-1}$;

  find a largest in modulus entry of Y_k called $V_{k,1} = (Y_k)_{i_1 j_1}$ in position (i_1, j_1);

  if $|(Y_k)_{i_1 j_1}| \leq 1 + \epsilon$ **then**

    return A_k;

  else

    for $s = 2, \ldots, r$ **do**

      find a largest in modulus entry of Y_k not in rows $i_1, \ldots, i_{s-1}$ or columns $j_1, \ldots, j_{s-1}$ called $(Y_k)_{i_s j_s}$ in position (i_s, j_s);

$$V_{k,s} = \begin{bmatrix} (Y_k)_{i_1 j_1} & (Y_k)_{i_1 j_2} & \cdots & (Y_k)_{i_1 j_s} \\ (Y_k)_{i_2 j_1} & (Y_k)_{i_2 j_2} & \cdots & (Y_k)_{i_2 j_s} \\ \vdots & \vdots & \ddots & \vdots \\ (Y_k)_{i_s j_1} & (Y_k)_{i_s j_2} & \cdots & (Y_k)_{i_s j_s} \end{bmatrix};$$

      if $\mathrm{vol}(V_{k,s}) \leq \mathrm{vol}(V_{k,s-1})$ **then**

        A_{k+1} equals A_k with row j_α replaced with row i_α in X for all $1 \leq \alpha \leq s-1$;

        exit loop;

      else if $s = r$ **then**

        A_{k+1} equals A_k with row j_α replaced with row i_α in X for all $1 \leq \alpha \leq r$;

Moreover, there are only finitely many submatrices in X, so the volume of the submatrices in X are bounded. □

Recall that we define

$$V_{k,s} = \begin{bmatrix} (Y_k)_{i_1 j_1} & (Y_k)_{i_1 j_2} & \cdots & (Y_k)_{i_1 j_s} \\ (Y_k)_{i_2 j_1} & (Y_k)_{i_2 j_2} & \cdots & (Y_k)_{i_2 j_s} \\ \vdots & \vdots & \ddots & \vdots \\ (Y_k)_{i_s j_1} & (Y_k)_{i_s j_2} & \cdots & (Y_k)_{i_s j_s} \end{bmatrix}$$

in Algorithm 12. There is a relationship between $\mathrm{vol}(V_{k,s})$ and $\mathrm{vol}(V_{k,s-1})$. Let

$$B_{k,s} = \begin{bmatrix} (Y_k)_{i_1 j_s} & (Y_k)_{i_2 j_s} & \cdots & (Y_k)_{i_{s-1} j_s} \end{bmatrix}^\top$$
$$C_{k,s} = \begin{bmatrix} (Y_k)_{i_s j_1} & (Y_k)_{i_s j_2} & \cdots & (Y_k)_{i_s j_{s-1}} \end{bmatrix}$$

Then

$$V_{k,s} = \begin{bmatrix} V_{k,s-1} & B_{k,s} \\ C_{k,s} & (Y_k)_{i_s j_s} \end{bmatrix}.$$

So by Lemma 65,

$$\mathrm{vol}(V_{k,s}) = \mathrm{vol}(V_{k,s-1}) \left| (Y_k)_{i_s j_s} - C_{k,s} V_{k,s-1}^{-1} B_{k,s} \right|. \tag{2.2}$$

By Equation (2.2) the condition that $\mathrm{vol}(V_{k,s}) > \mathrm{vol}(V_{k,s-1})$ is equivalent to the condition that

$$\left| (Y_k)_{i_s j_s} - C_{k,s} V_{k,s-1}^{-1} B_{k,s} \right| > 1, \tag{2.3}$$

in which case row j_s in A_k will be replaced with row i_s in X.

Note that Inequality 2.3 will hold if the sign of $(Y_k)_{i_s j_s}$ is different from the sign of $C_{k,s} V_{k,s-1}^{-1} B_{k,s}$ and $|(Y_k)_{i_s j_s}| > 1$. If we assume that the signs are chosen uniformly at random, then the probability that row j_s in A_k will be replaced with row i_s in X will be at least $\frac{1}{2}$ if $|(Y_k)_{i_s j_s}| > 1$.

Like the alternating maximum volume algorithm, Algorithm 10, we may generalize the greedy maximum volume algorithm to an *alternating greedy maximum volume algorithm* to find $r \times r$ dominant submatrices of $m \times n$ matrices.

We will not express the alternating greedy maximum volume algorithm in pseudocode due to notational complexity, but it is defined analogously to Algorithm 10 by alternating applying the iterations of Algorithm 12 between rows and columns.

We compare the average number of iterations taken to converge and the average processing time of Algorithm 8, Algorithm 11, and Algorithm 12. We fix r, then generate 50 random $2000 \times r$ matrices and initializations. We calculate the average number of iterations and average processing time needed to find dominant submatrices until the ratio of the volumes of successive submatrices is at most $1 + \epsilon$ where $\epsilon = 10^{-8}$ for various ranks. In Table 2.1, we give the average number of iterations taken to converge, and in Table 2.2, we give the average processing time taken to converge.

Table 2.1: Iterations to find a dominant submatrix of 50 random $2000 \times r$ matrices, $\epsilon = 10^{-8}$

	average # of iterations		
Rank r	Maximum Volume (Algorithm 8)	Simple Greedy Maximum Volume (Algorithm 11)	Greedy Maximum Volume (Algorithm 12)
100	61.84	39.76	30.74
200	80.70	53.18	40.44
300	95.62	63.36	44.70
400	100.14	69.16	49.60
500	109.30	73.28	53.40

Next, we compare the average number of iterations taken to converge and the average processing time for Algorithm 9, Algorithm 10, and an alternating version of Algorithm 12. By number of iterations, we mean the number of times a product of the form $A^{-1}X$ or XA^{-1} is calculated. We fix r, then generate 50 random 2000×2000 matrices and random initializations. We run the maximum volume algorithms until the ratio of the volumes of successive submatrices is at most $1 + \epsilon$ where $\epsilon = 10^{-8}$. In Table 2.3, we

Table 2.2: Time to find a dominant submatrix of 50 random $2000 \times r$ matrices, $\epsilon = 10^{-8}$

Rank r	average processing time (s): Maximum Volume (Algorithm 8)	Simple Greedy Maximum Volume (Algorithm 11)	Greedy Maximum Volume (Algorithm 12)
100	0.2342	0.1573	0.1254
200	0.7382	0.4965	0.3898
300	1.5782	1.0664	0.7606
400	2.5692	1.7747	1.3187
500	3.8953	2.6378	1.9743

give the average number of iterations taken to converge, and in Table 2.4, we give the average processing time taken to converge.

Table 2.3: Iterations to find a dominant submatrix of 50 random 2000×2000 matrices, $\epsilon = 10^{-8}$

Rank r	average # of iterations: 2D Maximum Volume (Algorithm 9)	Alternating Maximum Volume (Algorithm 10)	Alternating Greedy Maximum Volume (Alternating Algorithm 12)
100	132.56	76.82	41.30
200	187.72	102.60	50.96
300	215.20	122.14	64.60
400	233.48	132.56	65.20
500	240.28	142.88	70.22

Table 2.4: Time to find a dominant submatrix of 50 random 2000×2000 matrices, $\epsilon = 10^{-8}$

Rank r	average processing time (s): 2D Maximum Volume (Algorithm 9)	Alternating Maximum Volume (Algorithm 10)	Alternating Greedy Maximum Volume (Alternating Algorithm 12)
100	0.6476	0.3513	0.2055
200	2.0462	1.0892	0.5621
300	3.8986	2.1676	1.2013
400	6.7108	3.7724	1.9356
500	9.6167	5.6313	2.8996

As we can see from Table 2.3 and Table 2.4, the alternating greedy maximum volume algorithm performs better than the 2D maximum volume algorithm and the alternating maximum volume algorithm.

2.5 Maximum Volume Skeleton Decomposition

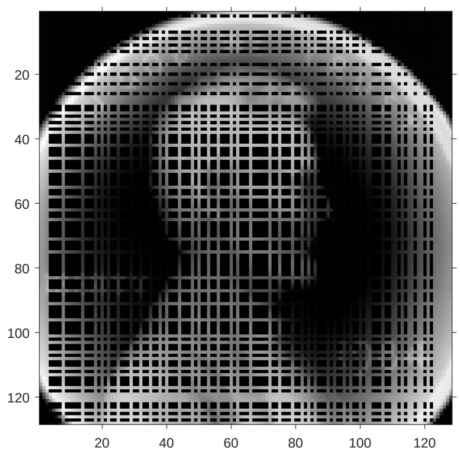
For our purposes, the primary application of maximum volume algorithms such as Algorithm 10 is to find an $r \times r$ dominant submatrix of an $m \times n$ matrix to use in the skeleton decomposition. We call such an application a *maximum volume skeleton decomposition*. Recall that the skeleton decomposition has least error when taken with respect to a maximum volume submatrix by Theorem 69. Also recall that dominant submatrices are not too far off in volume from maximum volume submatrices by Theorem 74.

In this section, we apply the maximum volume skeleton decomposition to compress various images with integer entries ranging from 0 to 255. We use the peak signal to noise ratio from Definition 10 to quantify the quality of the maximum volume skeleton decomposition.

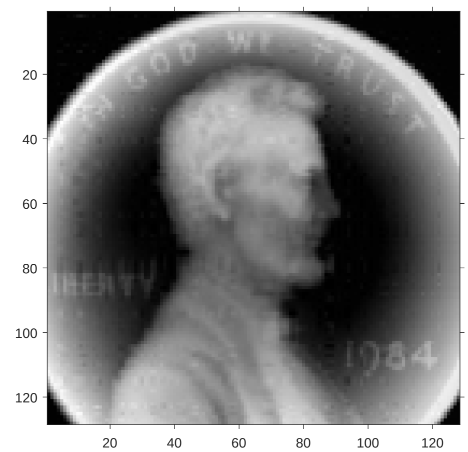
We first use the maximum volume skeleton decomposition to compress the 128×128 picture of a penny, Figure 1.5. Recall that each pixel corresponds to an integer from 0 to 255.

In Figure 2.1, we first choose a random 50×50 submatrix of Figure 1.5 which we use to initialize the alternating maximum volume algorithm, algorithm 10. Figure 2.1a shows the rows and columns corresponding to the output of the maximum volume algorithm, and Figure 2.1c shows the rows and columns corresponding to the original randomly chosen submatrix. Figure 2.1b shows the maximum volume skeleton decomposition, and Figure 2.1d shows the skeleton decomposition with respect to the initial randomly chosen submatrix. As can be seen, the maximum volume skeleton decomposition performs significantly better.

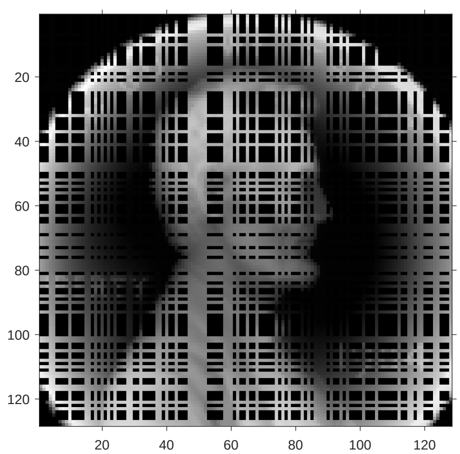
We apply the maximum volume skeleton decomposition to compress a variety of pictures. The maximum volume algorithm is initialized with a randomly chosen submatrix. We give the name of the image, the rank of the skeleton decomposition, peak signal to noise ratios, and the compression ratios in Table 2.5.



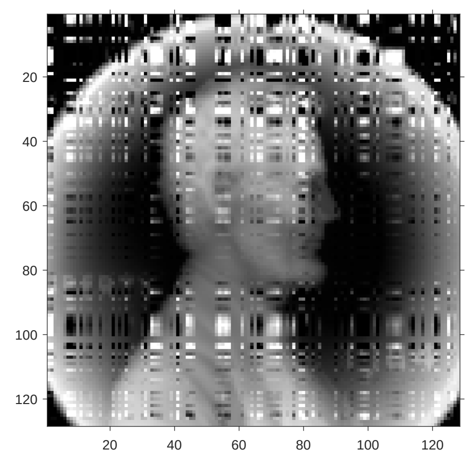
(a) 50 rows and 50 columns chosen by Algorithm 8



(b) Rank 50 skeleton decomposition with respect to submatrix chosen by maximum volume algorithm. PSNR $= 36.57$.



(c) 50 rows and 50 columns chosen randomly



(d) Rank 50 skeleton decomposition with respect to a randomly chosen submatrix. PSNR $= 10.01$.

Figure 2.1: A comparison of skeleton decompositions. The volume of the submatrix chosen by Algorithm 8 is larger than the volume of the randomly chosen submatrix by a factor of 10^{19}. Approximations have a compression ratio of 0.629.

Table 2.5: Maximum volume skeleton decomposition applied to various images

Image	Resolution	Rank	PSNR	Compression %
bang	512×512	350	32.23	0.900
Barbara	512×512	260	32.31	0.758
bike	512×512	420	32.18	0.968
boat	256×256	130	32.28	0.758
brain	512×512	115	33.53	0.400
clock	512×512	90	33.20	0.321
F16	512×512	230	32.22	0.697
finger	512×512	180	32.67	0.580
house	256×256	75	31.35	0.500
knee	512×512	105	32.21	0.368
monkey	512×512	440	32.38	0.980
MRI	512×512	105	32.73	0.368
pepper	512×512	375	32.61	0.928
Saturn	512×512	75	33.45	0.272

2.6 Johnson Graphs

In this section, we introduce Johnson graphs, which we relate to dominant submatrices. We use Johnson graphs to get an upper bound on the number of $r \times r$ dominant submatrices in almost all matrices.

Given a graph G, let $V(G)$ denote the set of vertices in G, and let $E(G)$ denote the set of edges in G.

Definition 81. *The* Johnson graph $J_{m,r}$ *is the graph whose vertices are the size* r *subsets of the* m *element set* $[m]$. *There is an edge connecting two vertices* I *and* I' *if their intersection contains exactly* r -1 *elements. That is,*

$$V(J_{m,r}) = \{I \subset [m] \mid |I| = r\}$$
$$E(J_{m,r}) = \{(I, I') \in V \times V \mid |I \cap I'| = r - 1\}.$$

The Johnson graph $J_{m,r}$ is an $r(m-r)$-regular graph, meaning each vertex has exactly $r(m-r)$ neighbors. $J_{m,r}$ has has $\binom{m}{r}$ vertices, $\frac{(m-r)r}{2}\binom{m}{r}$ edges, and has diameter $\min(r, m-r)$. Additionally, $J_{m,r} \cong J_{m,m-r}$.

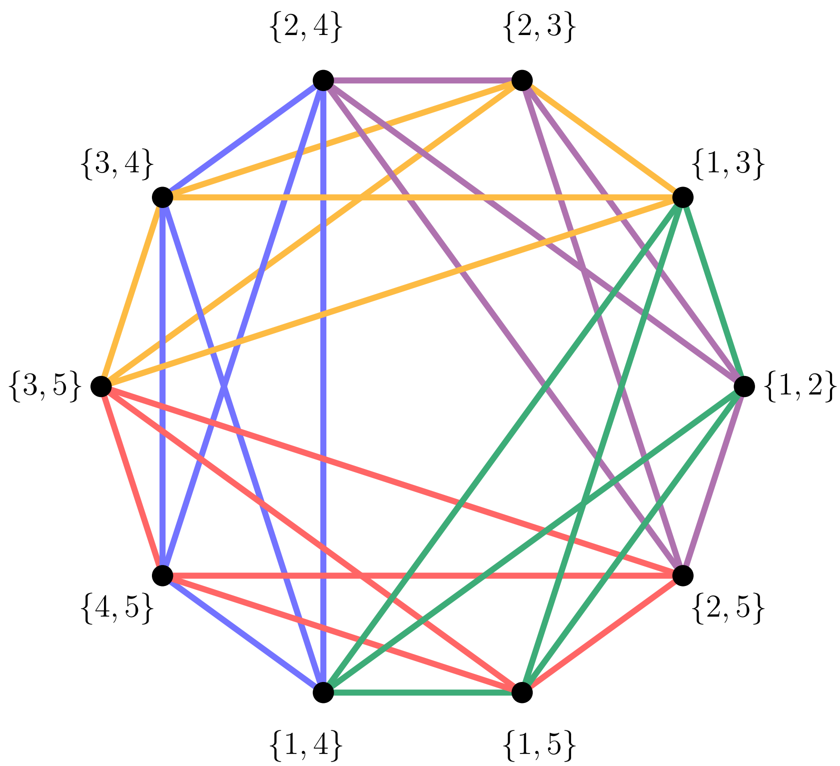


Figure 2.2: The Johnson graph $J_{5,2}$. Each vertex is a subset of $\{1, 2, 3, 4, 5\}$ of size two.

There are various ways to define the product of graphs. We introduce the most useful way for our purposes.

Definition 82. *Given graphs G and H, define the* graph Cartesian product *denoted $G \square H$ as the graph with vertices*

$$V(G \square H) = V(G) \times V(H).$$

Given vertices (u, v) and (u', v') in $V(G \square H)$, there is an edge between (u, v) and (u', v') in $G \square H$ if either

- *$u = u'$ and v is adjacent to v' in H, or*
- *$v = v'$ and u is adjacent to u' in G.*

Note that if H is the trivial graph with one vertex and no edges, then $G \square H \cong G$.

Given an $m \times n$ matrix X, define the map

$$F_X : J_{m,r} \Box J_{n,r} \to \mathbb{R}$$
$$(I, J) \mapsto \text{vol}(X_{I,J})$$

Although $X_{I,J}$ is only defined as a submatrix of X up to permutation of rows and columns, $\text{vol}(X_{I,J})$ is well defined because permuting rows or columns may only change the sign of the determinant.

Given a connected graph G, there is a distance metric on G where the distance between vertices u and u' is equal to the length of the shortest path between u and u'. Note that if a maximum volume submatrix $X_{\max}$ has entries in positions $I \times J$, then the vertex $(I, J) \in J_{m,r} \Box J_{n,r}$ is a global maximum of F_X.

It is in the following precise sense that we consider dominant submatrices to be submatrices with locally maximum volume.

Theorem 83. *Given an $m \times n$ matrix X, if an $r \times r$ dominant submatrix $\hat{X}$ has entries in positions $I \times J$ so $X_{I,J} = \hat{X}$, then the vertex $(I, J) \in J_{m,r} \Box J_{n,r}$ is a maximum of F_X in a closed ball of radius of one.*

Proof. Recall that an $r \times r$ submatrix $X_{I,J} = \hat{X}$ of X is considered dominant if the volume of $\hat{X}$ may not increase after swapping a row in $\hat{X}$ with a row in X in the same columns as $\hat{X}$, or by swapping a column in $\hat{X}$ with a column in X in the same rows as $\hat{X}$.

In other words, if $X_{I,J} = \hat{X}$ is dominant in X, then $\text{vol}(X_{I,J}) \geq \text{vol}(X_{I',J'})$ for all I' and J' such that

- $I = I'$ and $|J \cap J'| = r - 1$, or
- $J = J'$ and $|I \cap I'| = r - 1$.

Such (I', J') are exactly the vertices that are adjacent to (I, J) in $J_{m,r} \Box J_{n,r}$, or in other words, are the vertices in a closed ball of radius one centered at (I, J). □

The Johnson graph $J_{r,r}$ is trivial, so in the special case when $n = r$,

$$J_{m,r} \square J_{r,r} \cong J_{m,r}.$$

The maximum volume algorithm on $m \times r$ matrices, Algorithm 8, has the following interpretation. Given an $m \times r$ matrix X, on step k if $A_k = X_{I,[r]}$, we calculate which vertex $I' \in J_{m,r}$ adjacent to I will increase F_X by the largest amount by finding the largest in magnitude entry of XA_k^{-1}. We then let $A_{k+1} = X_{I',[r]}$, and repeat until a maximum of F_X in a closed ball of radius one is found.

More generally, the two directional maximum volume algorithm, Algorithm 9, has the following interpretation. Given an $m \times n$ matrix X, on step k, if $A_k = X_{I,J}$, we calculate which vertex $(I', J') \in J_{m,r} \square J_{n,r}$ adjacent to (I, J) will have the largest increase in F_X. We then let $A_{k+1} = X_{I',J'}$, and repeat until a maximum of F_X in a closed ball of radius one is found. In this sense, these maximum volume algorithms are discrete gradient ascent methods with step size equal to one.

There is often a unique maximum of a function because the only way a function can have more than one maximum is if there are two points of equal maximal value. For example, the only way that a matrix can have more than one maximum volume submatrix is if there are two submatrices of equal maximal volume. We use this principle along with Theorem 83 to obtain an upper bound on the number of dominant submatrices for almost all matrices.

Definition 84. *An* independent vertex set *in a graph G is a subset of the vertices of G such that no two vertices are adjacent in G. See Figure 2.3 for example. The* independence number *of a graph G is denoted* $\alpha(G)$ *and is defined as the maximum size of an independent vertex set in G.*

The only way that F_X can have more than one maximum point in a closed ball of radius one in $J_{m\times r} \square J_{n\times r}$ is if there are two dominant submatrices with positions in sets of indices which are adjacent in $J_{m,r} \square J_{n,r}$. This implies that there are two dominant submatrices of equal volume. In particular we have the following theorem.

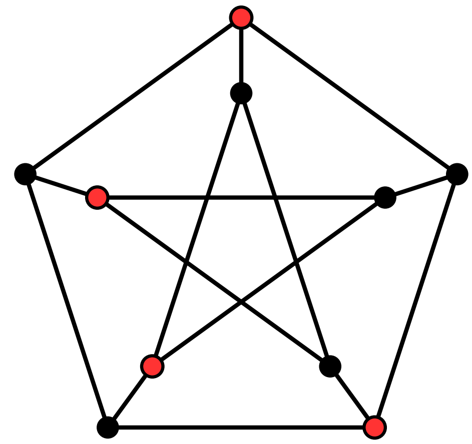


Figure 2.3: The set of red vertices is an independent vertex set.

Theorem 85. *Given an $m \times n$ matrix X, if no two $r \times r$ dominant submatrices have equal volume, then the independence number $\alpha(J_{m,r} \square J_{n,r})$ is an upper bound for the number of $r \times r$ dominant submatrices in X.*

Proof. Suppose the number of $r \times r$ dominant submatrices in X is larger than $\alpha(J_{m,r} \square J_{n,r})$. Then there are two dominant submatrices $X_{I,J}$ and $X_{I',J'}$ such that (I, J) is adjacent to (I', J') in $J_{m,r} \square J_{n,r}$, because otherwise the set of (I, J) such that $X_{I,J}$ is dominant would be an independent set in $J_{m,r} \square J_{n,r}$ with larger size than the independence number of $J_{m,r} \square J_{n,r}$. Because $X_{I,J}$ is dominant and (I, J) is adjacent to (I', J') in $J_{m,r} \square J_{n,r}$, then $\mathrm{vol}(X_{I,J}) \geq \mathrm{vol}(X_{I',J'})$ by Theorem 83. Similarly, because $X_{I',J'}$ is dominant and (I', J') is adjacent to (I, J) in $J_{m,r} \square J_{n,r}$, then $\mathrm{vol}(X_{I,J}) \leq \mathrm{vol}(X_{I',J'})$. So $\mathrm{vol}(X_{I,J}) = \mathrm{vol}(X_{I',J'})$. □

In other words, if the independence number $\alpha(J_{m,r} \square J_{n,r})$ is not an upper bound for the number of $r \times r$ dominant submatrices in a matrix X, then X must necessarily have least two dominant submatrices with the same volume. We show that the set of matrices that have at least two $r \times r$ submatrices with equal volume has measure zero in $M_{m \times n}$.

Theorem 86. *Let $V \subset M_{m \times n}$ be the set of matrices where each matrix has at least two distinct $r \times r$ submatrices of equal volume. Then V has measure zero in $M_{m \times n}$.*

Proof. Because

$$\mathrm{vol}(X) = |\det(X)| = \sqrt{\det(X)\overline{\det(X)}},$$

the condition that $\mathrm{vol}(X_{I,J}) = \mathrm{vol}(X_{I',J'})$ is equivalent to the condition that

$$\det(X_{I,J})\overline{\det(X_{I,J})} = \det(X_{I',J'})\overline{\det(X_{I',J'})}.$$

Here $\overline{z}$ denotes the complex conjugate of the complex number z. Define the set

$$V_{I,J,I',J'} = \{X \in M_{m\times n} \mid \det(X_{I,J})\overline{\det(X_{I,J})} - \det(X_{I',J'})\overline{\det(X_{I',J'})} = 0\}$$

where $I, I' \subset [m]$ and $J, J' \subset [n]$. Then

$$V = \bigcup_{I,J,I',J'} V_{I,J,I',J'}$$
$$\textit{s.t.}\ |I| = |J| = |I'| = |J'| = r$$
$$I \neq I' \text{ or } J \neq J'.$$

When $M_{m\times n} = \mathbb{R}^{m\times n}$ is the set of $m \times n$ real matrices,

$$V_{I,J,I',J'} = \{X \in M_{m\times n} \mid \det(X_{I,J})^2 - \det(X_{I',J'})^2 = 0\}.$$

$V_{I,J,I',J'}$ is the zero set of one non-trivial polynomial equation, and so has codimension at least one. Because V is a finite union of codimension at least one algebraic varieties, it has codimension at least one.

When $M_{m\times n} = \mathbb{C}^{m\times n}$ is the set of $m \times n$ complex matrices, note that

$$M_{m\times n} \cong \mathbb{C}^{mn} \cong \mathbb{R}^{2mn}.$$

If X is in $V_{I,J,I',J'}$, it must satisfy the non-trivial polynomial equations

$$\operatorname{Re}\left(\det(X_{I,J})\overline{\det(X_{I,J})} - \det(X_{I',J'})\overline{\det(X_{I',J'})}\right) = 0$$
$$\operatorname{Im}\left(\det(X_{I,J})\overline{\det(X_{I,J})} - \det(X_{I',J'})\overline{\det(X_{I',J'})}\right) = 0$$

where $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$ are the real and imaginary parts of the complex number z respectively. So $V_{I,J,I',J'}$ has codimension greater than zero in $\mathbb{R}^{2mn}$, which means V has codimension greater than zero.

Because the dimension of V is strictly smaller than the dimension of $M_{m\times n}$ in each case, V has measure zero in $M_{m\times n}$. □

Theorem 87. *The number of $r \times r$ dominant submatrices in an $m \times n$ matrix X up to permutation of rows and columns is at most $\alpha(J_{m,r}\Box J_{n,r})$ for almost all $X \in M_{m\times n}$.*

Proof. If $\alpha(J_{m,r}\Box J_{n,r})$ is not an upper bound for the number of $r \times r$ dominant submatrices in X, then X must have two submatrices of equal volume by Theorem 85. So X must be in the set V defined in Theorem 86 which has measure zero. □

We have established that $\alpha(J_{m,r}\Box J_{n,r})$ is an upper bound for number of $r \times r$ dominant submatrices in X for almost every $X \in M_{m\times n}$. Moreover, recall that $J_{r,r}$ is trivial, so $\alpha(J_{m,r})$ is the upper bound for the number of $r \times r$ submatrices of almost every $m \times r$ matrix.

There is no general formula for the independence number of the graph Cartesian product of graphs, but there is the following bound.

Theorem 88. *(Vizing, 1963) Given finite graphs G and H, the independence number satisfies*

$$\alpha(G\Box H) \leq \min\left(\alpha(G)\left|V(H)\right|, \alpha(H)\left|V(G)\right|\right).$$

So by Theorem 88, we have

$$\alpha(J_{m,r}\Box J_{n,r}) \leq \min\left(\alpha(J_{m,r})\binom{n}{r}, \alpha(J_{n,r})\binom{m}{r}\right). \tag{2.4}$$

The independence number of a Jonson graph $\alpha(J_{m,r})$ is not known in general. There is lots of literature on the independence number of Johnson graphs (Brouwer et al., 2006). It is known for example, that the independence number of the Johnson graph $J_{m,r}$ is equal to the size of the largest constant weight code of word length m, weight r, and distance at least four (Brouwer & Etzion, 2011). In other words, it is equal to the maximum number of binary vectors of length m having r ones and $m-r$ zeros such that any two vectors differ in at least four places.

There is an iterative upper bound on the independence number of Johnson graphs called the *Johnson bound*.

Theorem 89. *(Johnson, 1962) The Johnson bound states*

$$\alpha(J_{m,r}) \leq \frac{m}{r}\alpha(J_{m-1,r-1}).$$

So inductively we have

$$\alpha(J_{m,r}) \leq \frac{m!}{(m-r+1)!r!} = \frac{\binom{m}{r}}{1+m-r}. \tag{2.5}$$

Definition 90. *The* independence ratio *of a graph G is the ratio of the independence number of G to the number of vertices in G. That is, the independence ratio is equal to $\frac{\alpha(G)}{|V(G)|}$.*

Because the number of vertices in $J_{m,r}$ is equal to $\binom{m}{r}$, by Inequality (2.5), the independence ratio of $J_{m,r}$ satisfies the inequality

$$\frac{\alpha(J_{m,r})}{\binom{m}{r}} \leq \frac{1}{1+m-r}. \tag{2.6}$$

We get a similar bound for the independence ratio of $J_{m,r}\Box J_{n,r}$ using Inequality (2.4) and Inequality (2.6). Assuming $n \leq m$, we have the bound

$$\frac{\alpha(J_{m,r}\Box J_{n,r})}{\binom{m}{r}\binom{n}{r}} \leq \frac{1}{1+m-r}.$$

The independence number of Johnson graphs is known in a few cases. For example

$$\begin{aligned}\alpha(J_{m,1}) &= 1\\ \alpha(J_{m,m-1}) &= 1\\ \alpha(J_{m,2}) &= \left\lfloor \frac{m}{2} \right\rfloor\\ \alpha(J_{m,1}\Box J_{n,1}) &= \min(m,n)\\ \alpha(J_{m,m-1}\Box J_{n,n-1}) &= \min(m,n).\end{aligned}$$

It is fairly straightforward to find examples of matrices that have these numbers of dominant submatrices. To show the upper bound provided in Theorem 85 is sharp the examples must also have the property that no two dominant submatrices have entries with positions in indices which are adjacent in $J_{m,r}\Box J_{n,r}$. We construct an explicit example in the case when $n = r = 2$,

Example 91. *We show that* $\alpha(J_{m,2}) = \lfloor \frac{m}{2} \rfloor$ *is a sharp upper bound for the number of* 2×2 *dominant submatrices for almost all* $m \times 2$ *matrices by constructing an* $m \times 2$ *matrix such that no two dominant submatrices have entries with positions in sets of indices which are adjacent in* $J_{m,2}$.

First let us assume that m *is even. Let*

$$A_i = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix},$$

where $\theta_i = \frac{\pi(i-1)}{m}$ *for* $i = 1, \ldots, \frac{m}{2}$. *We show that each of* A_i *are the dominant submatrices of the matrix*

$$X = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{\frac{m}{2}} \end{bmatrix}.$$

Moreover, the submatrix A_i *has entries in positions* $\{2i-1, 2i\} \times \{1, 2\}$.

Note that

$$A_i^{-1} = \begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{bmatrix}$$

$$A_j A_i^{-1} = \begin{bmatrix} \cos(\theta_j - \theta_i) & -\sin(\theta_j - \theta_i) \\ \sin(\theta_j - \theta_i) & \cos(\theta_j - \theta_i) \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\frac{\pi(j-i)}{m}) & -\sin(\frac{\pi(j-i)}{m}) \\ \sin(\frac{\pi(j-i)}{m}) & \cos(\frac{\pi(j-i)}{m}) \end{bmatrix}.$$

Because $0 \leq \theta_i < \frac{\pi}{2}$ *we have* $-\frac{\pi}{2} < \theta_j - \theta_i < \frac{\pi}{2}$, *and so* $|\sin(\theta_j - \theta_i)| < 1$. *Moreover,* $\theta_j - \theta_i = 0$ *if and only if* $j = i$, *and so* $|\cos(\theta_j - \theta_i)| = 1$ *if and only if* $j = i$. *So* A_i *are dominant submatrices for all* i, *and because each entry of* $A_j A_i^{-1}$ *is strictly less than one in modulus when* $j \neq i$, *this implies that there is no index* I *adjacent to* $\{2i-1, 2i\}$ *in* $J_{m,2}$ *such that* $X_{I,[2]}$ *is dominant by Theorem 83.*

When m *is odd the* $m \times 2$ *matrix*

$$X = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{\lfloor \frac{m}{2} \rfloor} \\ 0 \end{bmatrix}$$

similarly has $\lfloor \frac{m}{2} \rfloor$ *dominant submatrices, no two of which are adjacent.*

Theorem 87 says that for almost all $X \in M_{m \times n}$, $\alpha(J_{m,r} \square J_{n,r})$ is an upper bound for the number of $r \times r$ dominant submatrices in X. We conjecture that the upper bound is sharp on a set of positive measure.

Conjecture 92. *There is a subset $U \subset M_{m\times n}$ of positive measure such that $\alpha(J_{m,r}\square J_{n,r})$ is equal to the number of $r \times r$ dominant submatrices in X for all $X \in U$. In other words, $\alpha(J_{m,r}\square J_{n,r})$ is the essential supremum of the function which counts the number of $r \times r$ dominant submatrices.*

We give the following theorem to support Conjecture 92.

Theorem 93. *Let X be an $m \times n$ matrix such that no two dominant $r \times r$ submatrices have row and column index sets which are adjacent in $J_{m,r}\square J_{n,r}$. Then there exists $\epsilon > 0$ such that for all Y in the ball $B_\epsilon(X)$, the dominant $r \times r$ submatrices in Y have the same positions as the dominant $r \times r$ submatrices in X, and so X and Y have the same number of dominant submatrices.*

Proof. Given given an $r \times r$ submatrix A of X, first suppose A is dominant. After permuting rows and columns, suppose X has the block structure $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ for corresponding submatrices B, C, and D. Because no two dominant submatrices have index sets which are adjacent in $J_{m,r}\square J_{n,r}$, another dominant submatrix may not be obtained from A by swapping one pair of rows or columns by Theorem 83. Therefore, each entry of BA^{-1} and $A^{-1}C$ is strictly smaller than one in modulus by Lemma 70. Define the functions

$$f(X) = \left\|BA^{-1}\right\|_\infty$$
$$g(X) = \left\|A^{-1}C\right\|_\infty .$$

Then $f(X) < 1$ and $g(X) < 1$. Because f and g are continuous, there exists $\epsilon_A > 0$ such that for all $Y \in B_{\epsilon_A}(X)$, $f(Y) < 1$ and $g(Y) < 1$, so the submatrix of Y in the same position as A in X is dominant in Y.

Now suppose A is not a dominant submatrix of X. Then because A is not dominant, $f(X) > 1$ and $g(X) > 1$. Similarly, because f and g are continuous, there exists $\epsilon_A > 0$ such that for all $Y \in B_{\epsilon_A}(X)$, $f(Y) > 1$ and $g(Y) > 1$, so the submatrix of Y in the same position as A in X is not dominant in Y.

Because there are only finitely many $r \times r$ submatrices, let

$$\epsilon = \min_A \{\epsilon_A\}.$$

Then for all $Y \in B_\epsilon(X)$, the dominant submatrices in Y have the same positions as the dominant submatrices in X. So the number of dominant submatrices in Y is equal to the number of dominant submatrices in X. $\square$

Theorem 93 implies that, if there exists at least one $m \times n$ matrix X with $\alpha(J_{m,r} \square J_{n,r})$ dominant submatrices such that no two dominant $r \times r$ submatrices have row and column index sets which are adjacent in $J_{m,r} \square J_{n,r}$, then Conjecture 92 is true. This is because the set $B_\epsilon(X)$ would be a positive measure set of matrices such that each matrix has $\alpha(J_{m,r} \square J_{n,r})$ dominant $r \times r$ submatrices. If the volume of each submatrix could be chosen independently, or even if only the ordering of the volumes of the submatrices could be chosen, then such a matrix would exist. For example, let $S \subset V(J_{m,r} \square J_{n,r})$ be an independent vertex set of maximal size. Suppose a matrix $X \in M_{m \times n}$ has the property that

$$\text{vol}(X_{I,J}) = 1 \text{ if } (I, J) \in S$$
$$\text{vol}(X_{I,J}) = 0 \text{ if } (I, J) \notin S.$$

Then the number of $r \times r$ dominant submatrices in X is equal to $\alpha(J_{m,r} \square J_{n,r})$, and no two dominant $r \times r$ submatrices have row and column index sets which are adjacent in $J_{m,r} \square J_{n,r}$. The complication is that for $r > 1$, the volumes of submatrices are interdependent, and cannot be chosen individually.

In general, there will be fewer than $\alpha(J_{m,r} \square J_{n,r})$ dominant submatrices.

Example 94. *Consider the 1×1 submatrices of a random 2×2 matrix $X = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ where each entry is sampled independently from an identical continuous density. Assume that no two entries are equal, which happens with probability one. By permuting rows and columns, assume a is the largest entry in modulus.*

Then there are $3! = 6$ *possibilities for the order of the entries*

$$|a| > |b| > |c| > |d|$$
$$|a| > |c| > |b| > |d|$$
$$|a| > |b| > |d| > |c|$$
$$|a| > |c| > |d| > |b|$$
$$|a| > |d| > |b| > |c|$$
$$|a| > |d| > |c| > |b|$$

each of which happen with probability $1/6$*. In the first through fourth cases,* a *is the only dominant submatrix. In cases five and six, both* a *and* d *are dominant submatrices. So the probability that* X *will have one* 1×1 *dominant submatrix is* $2/3$*, and the probability that* X *will have two* 1×1 *dominant submatrices is* $1/3$*.*

2.7 Maximum Volume Gradient Descent for Matrix Completion

Many existing methods for matrix completion such as alternating projection, Algorithm 1, and singular value thresholding, Algorithm 4, calculate the singular value decomposition of a matrix at each step, which is computationally expensive.

It is worth noting that orthogonal rank one matrix pursuit, Algorithm 3, only calculates the first singular value and singular vectors at each step, and alternating minimization, Algorithm 2 does not use the singular value decomposition.

We seek a matrix completion method that is efficient, accurate, and scalable. In this section we introduce a maximum volume Schur complement based gradient descent method for low-rank matrix completion. One nice property of this method is that the gradient of the objective function is a rational function of the actual entries of the matrix. This means that it is not necessary to compute a singular value decomposition at each iteration as is needed in

Recall Example 32 where the $m \times n$ partially known matrix M_Ω has the structure

$$M_\Omega = \begin{bmatrix} A & B \\ C & \square \end{bmatrix}$$

where A, B, and C are fully known submatrices, A is $r \times r$ nonsingular, B is $r \times (n-r)$, C is $(m-r) \times r$, and $\square$ is an $(m-r) \times (n-r)$ block of unknown entries. M_Ω has the unique rank r completion

$$M = \begin{bmatrix} A & B \\ C & CA^{-1}B \end{bmatrix}.$$

An arbitrary partially known matrix M_Ω is unlikely to have this structure. However, if M is a rank r completion of M_Ω, then the rows and columns of M may be permuted so that $M = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ where A is $r \times r$ nonsingular and $D - CA^{-1}B = 0$.

To recover missing entries from M_Ω, we may cast the low-rank matrix completion problem as the minimization of the norm of a Schur complement

$$\begin{aligned} &\min \frac{1}{2} \left\| D - CA^{-1}B \right\|^2 \\ &\textit{s.t. } P_\Omega\left(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]\right) = M_\Omega. \end{aligned} \tag{2.7}$$

Note that $\frac{1}{2} \left\| D - CA^{-1}B \right\|^2 \geq 0$, and $\frac{1}{2} \left\| D - CA^{-1}B \right\|^2 = 0$ if and only if $\operatorname{rank}(\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]) = r$ by Lemma 66. So if M_Ω has a rank r completion $M = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ such that A is nonsingular, then M is a solution to Minimization (2.7). In general, A may be any $r \times r$ nonsingular submatrix.

We solve Minimization (2.7) using gradient decent. Let $\langle X, Y \rangle = \operatorname{tr}(X^\top Y)$ denote the dot product of the matrices X and Y. Then $\|X\|^2 = \langle X, X \rangle$. We first introduce the following identities from matrix calculus.

Proposition 95. *(Petersen, Pedersen, et al., 2008) Consider a function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$. Let $\frac{\partial f}{\partial X}$ denote the $m \times n$ matrix where $(\frac{\partial f}{\partial X})_{ij} = \frac{\partial f}{\partial x_{ij}}$. Then for a function $F : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$,*

$$\begin{aligned} \partial f(X) &= \langle F(X), \partial X\rangle \\ \implies \frac{\partial f(X)}{\partial X} &= F(X). \end{aligned}$$

We also have the identities

$$\begin{aligned} \partial(X^\top) &= (\partial X)^\top \\ \partial(\operatorname{tr}(X)) &= \operatorname{tr}(\partial X) \\ \partial(XY) &= (\partial X)Y + X(\partial Y) \\ \partial\left(\frac{1}{2}\|X\|^2\right) &= \langle X, \partial X\rangle \\ \partial(X^{-1}) &= -X^{-1}(\partial X)X^{-1}. \end{aligned}$$

Let $X = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ be the $m \times n$ matrix where A is $r \times r$, B is $r \times (n-r)$, C is $(m-r) \times r$, and D is $(m-r) \times (n-r)$. Let $S_A = D - CA^{-1}B$ denote the Schur complement of X with respect to A. Define the loss function

$$\begin{aligned} f(X) &= \frac{1}{2}\left\|D - CA^{-1}B\right\|^2 \\ &= \frac{1}{2}\|S_A\|^2 \end{aligned}$$

We use Proposition 95 to calculate the derivatives of f.

Theorem 96. *We have the gradient information*

$$\frac{\partial f}{\partial A} = A^{-\top}C^{\top}S_AB^{\top}A^{-\top}$$
$$\frac{\partial f}{\partial B} = -A^{-\top}C^{\top}S_A$$
$$\frac{\partial f}{\partial C} = -S_AB^{\top}A^{-\top}$$
$$\frac{\partial f}{\partial D} = S_A$$

Proof. Using identities from Proposition 95, we have

$$\begin{aligned}
\partial f &= \langle S_A, \partial S_A\rangle \\
&= \langle S_A, \partial(D - CA^{-1}B)\rangle \\
&= \langle S_A, \partial D\rangle - \langle S_A, \partial(CA^{-1}B)\rangle \\
&= \langle S_A, \partial D\rangle - \langle S_A, (\partial C)A^{-1}B\rangle - \langle S_A, C(\partial A^{-1})B\rangle - \langle S_A, CA^{-1}(dB)\rangle \\
&= \langle S_A, \partial D\rangle - \langle S_AB^{\top}A^{-\top}, \partial C\rangle + \langle C^{\top}S_AB^{\top}, A^{-1}(\partial A)A^{-1}\rangle - \langle A^{-\top}C^{\top}S_A, \partial B\rangle \\
&= \langle S_A, \partial D\rangle - \langle S_AB^{\top}A^{-\top}, \partial C\rangle + \langle A^{-\top}C^{\top}S_AB^{\top}A^{-\top}, \partial A\rangle - \langle A^{-\top}C^{\top}S_A, \partial B\rangle.
\end{aligned}$$

Every term but the term with ∂Z vanishes when calculating $\frac{\partial f}{\partial Z}$ where $Z = A, B, C$, or D, so we have the desired result. □

Assembling the derivative information from Theorem 96, we have we have

$$\begin{aligned}
\frac{\partial f}{\partial X} &= \begin{bmatrix} \frac{\partial f}{\partial A} & \frac{\partial f}{\partial B} \\ \frac{\partial f}{\partial C} & \frac{\partial f}{\partial D} \end{bmatrix} \\
&= \begin{bmatrix} A^{-\top}C^{\top}S_AB^{\top}A^{-\top} & -A^{-\top}C^{\top}S_A \\ -S_AB^{\top}A^{-\top} & S_A \end{bmatrix}.
\end{aligned}$$

Note that entries are rational functions of the elements of X. We now present a Schur complement based gradient decent method, Algorithm 13.

Algorithm 13: Schur Complement Gradient Descent

Input: partially known matrix M_Ω, initial guess $X_0 = \left[\begin{smallmatrix} A_0 & B_0 \\ C_0 & D_0 \end{smallmatrix}\right]$ such that $P_\Omega(X_0) = M_\Omega$, sequence of step sizes $\{h_k\}$, stopping criterion

Result: $X_k = \left[\begin{smallmatrix} A_k & B_k \\ C_k & D_k \end{smallmatrix}\right]$ an approximation of a rank r completion of M_Ω

for $k = 0, 1, \ldots$ **do**

$S_{A_k} = D_k - C_k(A_k^{-1})B_k$;

$\left[\begin{smallmatrix} A_{k+1} & B_{k+1} \\ C_{k+1} & D_{k+1} \end{smallmatrix}\right] = \left[\begin{smallmatrix} A_k & B_k \\ C_k & D_k \end{smallmatrix}\right] - h_k P_{\Omega^c}\left(\left[\begin{smallmatrix} A_k^{-\top} C_k^\top S_{A_k} B_k^\top A_k^{-\top} & -A_k^{-\top} C_k^\top S_{A_k} \\ -S_{A_k} B_k^\top A_k^{-\top} & S_{A_k} \end{smallmatrix}\right]\right)$;

It is undesirable to change the known entries of X_k in Algorithm 13. The operator P_{Ω^c} is applied to the gradient which sets all entries with positions in Ω equal to zero, this way $P_\Omega(X_k) = M_\Omega$ for all k under the condition that $P_\Omega(X_0) = M_\Omega$. Empirically, $\sigma_{r+1}(X_k)$ is decreasing in k. However, $\sigma_1(X_k)$ tends to increase and diverge, so improvements must be made.

To improve the stability and convergence of Algorithm 13, We alternate between taking gradient descent steps and applying a maximum volume algorithm. Recall that $\|D - CA^{-1}B\|_{\max}$ is minimized when A is chosen with maximum volume by Theorem 69. Moreover, if S_A is $(m-r) \times (n-r)$, then

$$\|S_A\| \leq \sqrt{(m-r)(n-r)}\,\|S_A\|_{\max},$$

So minimizing $\|S_A\|_{\max}$ assists in solving Minimization (2.7). We introduce the Schur complement based maximum volume gradient descent, Algorithm 14.

Note that the gradient in Algorithm 14 may be of a different function at each iteration. In particular, given $I \subset [m]$ and $J \subset [n]$ such that $|I| = |J| = r$. In particular, let

$$f_{I,J}(X) = \left\| X_{I^c,J^c} - X_{I^c,J}(X_{I,J}^{-1})X_{I,J^c} \right\|.$$

Then on step k, the gradient is of f_{I_k,J_k}. However, note that $f_{I,J}(X) = 0$ if and only if $X_{I,J}$ is invertible and $\operatorname{rank}(X) = r$ for all I and J.

Algorithm 14: Maximum Volume Gradient Descent for Matrix Completion

Input: partially known matrix M_Ω, initial guess X_0 such that $P_\Omega(X_0) = M_\Omega$, $r \times r$ nonsingular submatrix $A_0 = (X_0)_{I_0,J_0}$ of X_0, sequence of step sizes $\{h_k\}$, tolerance $\epsilon > 0$, stopping criterion

Result: $X_k = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}$ an approximation of a rank r completion of M_Ω

for $k = 1, \dots$ **do**

$A_k = (X_{k-1})_{I_k,J_k}$, $I_k \subset [m]$, $J_k \subset [n]$ is the output of a two directional maximum volume algorithm, such as Algorithm 10 or an alternating version of Algorithm 12, with initialization X_{k-1}, initial nonsingular submatrix $(X_{k-1})_{I_{k-1},J_{k-1}}$, and tolerance ϵ;

find I_k^c and J_k^c, the complements of I_k and J_k;

$B_k = (X_k)_{I_k,J_k^c}$;

$C_k = (X_k)_{I_k^c,J_k}$;

$D_k = (X_k)_{I_k^c,J_k^c}$;

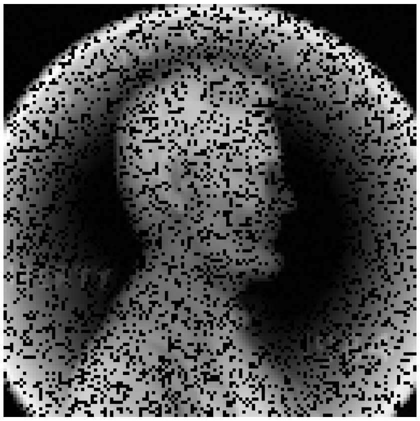
$S_{A_k} = D_k - C_k(A_k^{-1})B_k$;

$X_k = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} - h_k P_{\Omega^c}\left(\begin{bmatrix} A_k^{-\top} C_k^\top S_{A_k} B_k^\top A_k^{-\top} & -A_k^{-\top} C_k^\top S_{A_k} \\ -S_{A_k} B_k^\top A_k^{-\top} & S_{A_k} \end{bmatrix}\right)$;

Example 97. *Like Example 12, we delete* 25% *of the* 128×128 *picture of the United States penny, Figure 1.5, which we recover with Algorithm 14. For the initial guess, we first take the known entries of the image with unknown entries set to zero. We then apply three iterations of the alternating projection algorithm, Algorithm 1 to improve the initial guess. Also like Example 12, we set entries less than* 0 *or greater than* 255 *equal to* 0 *and* 255 *respectively at each iteration.*

Again, we can see in Figure 2.5 that the singular values of Figure 1.5 taper off at around $r = 18$*, and the singular values of the recovered picture closely matches the singular values of the original picture.*

In general it is difficult to prove the convergence of Algorithm 13 and Algorithm 14 because Minimization (2.7) is not convex. Starting with any initial matrix may not converge to a minimum, a good initial guess is important for convergence. One may improve the initial guess by taking a few initial iterations with methods such as Algorithm 1. Alternatively, one may replace the low-rank approximation step in Algorithm 1 with a maximum volume skeleton decomposition step.



(a) 25% of Figure 1.5 deleted uniformly at random

(b) Recovered image from Figure 2.4a with Algorithm 14

Figure 2.4: Figure 1.5 recovered with 50,000 iterations of Algorithm 14 assuming the rank of the original image is 18. Step size $h = 10^{-3}$, $\epsilon = 10^{-8}$. Initial 18×18 nonsingular submatrix chosen randomly. Recovery has a peak signal to noise ratio of 31.71.

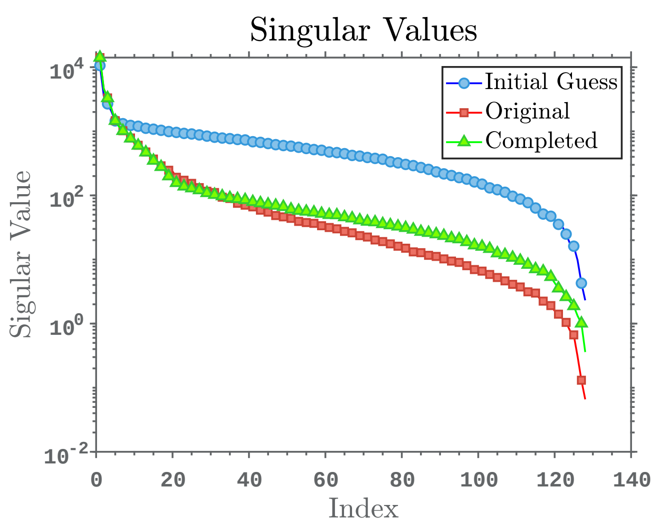


Figure 2.5: Singular values of Figure 1.5, Figure 2.4a, Figure 2.4b

Chapter 3

Maximum Volume Skeletal Decomposition for Scalable Plasma Physics Applications

3.1 Introduction

Oak Ridge National Lab (ORNL) is leading several branches of research effort to enable the successful design and operation of fusion reactor scale devices including the international effort at the International Thermonuclear Experimental Reactor (ITER) and the national effort towards a fusion pilot plant. To date, 16 megawatts of energy have successfully been produced in the large magnetic confinement device Joint European Torus (JET). ITER has been designed to generate 1500 megawatts of energy, which is orders of magnitude greater (Federici et al., 2001).

ITER is based on the tokamak, a toroidal device that confines extremely hot plasma with an axisymmetric magnetic field and is presently the most promising path to the production of fusion power. The plasma at the edge of confinement acts as a buffer between the high temperatures in the plasma core and the escaping heat flux that gets deposited on the plasma facing components of the device.

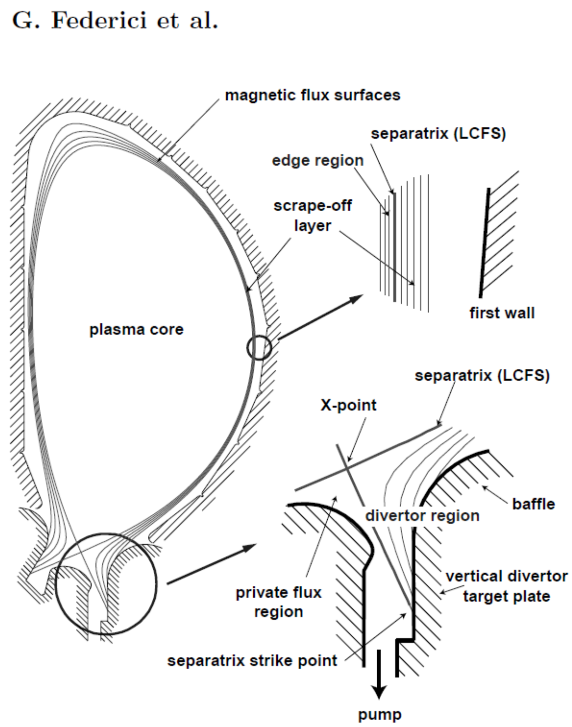


Figure 3.1: Schematic of poloidal cross-section of the ITER tokamak device (Federici et al., 2001)

See for example Figure 3.1 and Figure 3.3, which shows an illustration of the tokamak scheme used to contain a plasma at temperatures hotter than the surface of the sun. At high enough density and temperature, the plasma in the core of the device will undergo fusion with the proper fuel. The fusion produces energetic neutrons and energy sufficient to sustain the reaction. Turbulence carries some of the particles across magnetic fields from closed flux surfaces to open flux surfaces that intersect the device walls. At the scrape-off layer, plasma is directed along magnetic field lines down to the divertor component which is designed to handle these excessive heat flux loads and preserve the first wall. The separatrix indicates the boundary between these two regions of a tokamak plasma. Understanding the dynamics of the edge plasma is crucial to protect the material of the device and allow for operation at the reactor scale.

The Scrape-Off Layer Plasma Simulation (SOLPS-ITER) code package is the state-of-the-art tool used to predict the steady-state physics of this tokamak boundary plasma (Bonnin et al., 2016). SOLPS-ITER solves the multi-fluid plasma equations for continuity, momentum, and energy with a kinetic Monte Carlo description for the plasma interactions with the neutral gas used to control and dissipate the heat flux on the divertor. For a reactor configuration, SOLPS-ITER is estimated to take upwards of weeks-months of compute time in order to converge upon a valid solution (Kaveeva et al., 2018).

3.2 Dynamic Mode Decomposition

Researchers within the Fusion Energy Division at ORNL aim to reduce the computational complexity of fusion plasma simulation via construction of linear time advance operators for multi-timescale problems. This work is based on the data-driven dynamic mode decomposition (DMD) method, which has proven to be a useful tool for the analysis of fluid dynamics (Tu, 2013). The DMD relies on the singular value decomposition (SVD). However, for high dimensional data the SVD scales poorly, on the order of $O(m^2n)$ for $m \times n$ matrices with $m \leq n$. The matrices to be approximated are generated by the finite volume SOLPS-ITER code for modeling the flow of plasma at the edge region of a tokamak fusion device (Bonnin et al., 2016). A typical $m \times n$ SOLPS-ITER run produces a data matrix that can have m on the order of 1000 for spatial coordinates, and n greater than $10,000$ in the number of time steps over the roughly ten computed variables. For kinetic plasma simulations and magnetohydrodynamic plasma simulations of the entire core these sizes may be orders of magnitude greater. Application of the SVD to the high dimensional data required for fusion simulation is computationally prohibitive and alternative methods are desirable, as has been investigate in (Hatch et al., 2012) and (del-Castillo-Negrete et al., 2007). We aim to leverage the skeleton decomposition defined in Definition 68 with respect to a submatrix chosen by the alternating maximum volume algorithm presented in Algorithm 10 or an alternating version of the greedy maximum volume algorithm presented in Algorithm 12 to allow for the extension of the DMD to the scales of simulation data inherent in the physics of fusion plasmas.

We first introduce the Moore-Penrose pseudoinverse.

Definition 98. *The* pseudoinverse *of a matrix X is denoted $X^\dagger$ and is defined as follows. Let r be the rank of X. Compute the SVD $X = U\Sigma V^*$, where Σ is the $r \times r$ diagonal matrix of non-zero singular values, U is the $m \times r$ matrix of r left-singular vectors, and V is the $n \times r$ matrix of r right-singular vectors. The* *pseudoinverse of X is defined as $X^\dagger = V\Sigma^{-1}U^*$. Let P_X denote the orthogonal projection operator onto the column space of X. Then*

$$P_X = XX^\dagger.$$

We now present the DMD algorithm defined in (Tu, 2013). Consider a sequence of real valued data vectors $\{z_0, \ldots, z_n\}$ where $z_k \in \mathbb{R}^m$ for all k. Suppose the vectors satisfy the linear relationship $z_{k+1} = Az_k$ for some $m \times m$ matrix A. Define the $m \times n$ matrices $X = [z_0 \ \cdots \ z_{n-1}]$ and $Y = [z_1 \ \cdots \ z_n]$. Then X and Y satisfy the equation

$$Y = AX.$$

Define $A = YX^\dagger$. If the vectors of x_k are linearly independent, then the equation $AX = Y$ has the exact solution $A = YX^\dagger$ which minimizes $\|A\|$. If there is no exact solution, then $A = YX^\dagger$ minimizes the quantity $\|AX - Y\|$ (Tu, 2013).

Compute the SVD of X obtaining $X = U\Sigma V^*$ Find a closest rank r approximation X_r of X using the first r singular values of X and the first r columns of U and V. By Theorem 9,

$$X_r = U_r \Sigma_r V_r^* \tag{3.1}$$

where U_r is the $m \times r$ matrix consisting of the first r left-singular vectors, V_r is $n \times r$ matrix consisting of the first r right-singular vectors, and Σ_r is the $r \times r$ diagonal matrix consisting of the first r singular values of X.

Define the matrices

$$A_r = YX_r^\dagger = YV_r\Sigma_r^{-1}U_r^* \tag{3.2}$$

$$\tilde{A}_r = U_r^* A_r U_r = U_r^* Y V_r \Sigma_r^{-1}. \tag{3.3}$$

It is assumed that A_r will be a suitable substitute for A when used to calculate the dynamics of the system. We can find $\tilde{z}_{n+N}$, an approximation of the unobserved data vector z_{n+N} for $N > 0$ by calculating

$$
\begin{aligned}
\tilde{z}_{n+N} &= (P_{X_r} A_r)^N P_{X_r} z_n \\
&= (U_r U_r^* A_r)^N (U_r U_r^*) z_n \\
&= U_r \tilde{A}_r^N U_r^* z_n.
\end{aligned}
$$

We use the fact that

$$
\begin{aligned}
P_{X_r} &= X_r X_r^\dagger \\
&= U_r \Sigma_r V_r^* V_r \Sigma_r^{-1} U_r^* \\
&= U_r U_r^*.
\end{aligned}
$$

Note that $V_r^* V_r = I_r$, where I_r is the $r \times r$ identity matrix, because V_r has non-zero orthogonal columns.

Because of the long processing time required to compute a low-rank approximation of X with the singular value decomposition, an alternative low-rank approximation technique is desirable. We use the skeleton decomposition as a faster alternative at the cost of a larger error in the low-rank approximation.

The skeleton decomposition of X is taken with respect to an $r \times r$ non-singular submatrix $\hat{X}$ in corresponding columns C and rows R. C is $m \times r$ and R is $r \times n$. The skeleton decomposition of X with respect to $\hat{X}$ is

$$
X_r = C \hat{X}^{-1} R \tag{3.4}
$$

where $\text{rank}(X_r) = r$. Because $\hat{X}$ is full rank, R and C are also full rank, so

$$
\begin{aligned}
X_r^\dagger &= R^\dagger \hat{X} C^\dagger \\
C^\dagger &= (C^* C)^{-1} C^* \\
R^\dagger &= R^* (R R^*)^{-1}.
\end{aligned}
$$

We show that several theorems from (Tu, 2013) also hold when the low-rank approximation of X is taken with respect to the skeleton decomposition. Like Equation (3.2) and Equation (3.3), define the matrices

$$A_r = YX_r^\dagger = YR^\dagger \hat{X} C^\dagger \tag{3.5}$$

$$\tilde{A}_r = C^\dagger Y R^\dagger \hat{X} = C^\dagger A_r C \tag{3.6}$$

Note that $C^\dagger C = I_r$ because the columns of C are linearly independent.

We consider the eigenvalues λ and corresponding eigenvectors w of $\tilde{A}_r$, so

$$\tilde{A}_r w = \lambda w.$$

Definition 99. *The* projected skeletal DMD mode *with respect to the eigenvalue λ is given by*

$$\hat{\varphi} = Cw. \tag{3.7}$$

The exact skeletal DMD mode with respect to λ is given by

$$\varphi = \frac{1}{\lambda} Y R^\dagger \hat{X} w. \tag{3.8}$$

Define the matrix $B = YR^\dagger \hat{X}$. Then in terms of B, we have

$$\begin{aligned} A_r &= BC^\dagger \\ \tilde{A}_r &= C^\dagger B \\ \varphi &= \frac{1}{\lambda} Bw. \end{aligned}$$

Theorem 100. *φ is an eigenvector of A_r with eigenvalue λ. Moreover, every non-zero eigenvalue of A_r is also an eigenvalue of $\tilde{A}_r$.*

Proof. We have

$$\begin{aligned}
A_r\varphi &= (BC^\dagger)(\frac{1}{\lambda}Bw) \\
&= B(\frac{1}{\lambda}\tilde{A}_r w) \\
&= Bw \\
&= \lambda(\frac{1}{\lambda}Bw) \\
&= \lambda\varphi
\end{aligned}$$

Moreover, $\varphi \neq 0$ because if $\varphi = \frac{1}{\lambda}Bw = 0$, then $C^\dagger Bw = \tilde{A}_r w = \lambda w = 0$, so $\lambda = 0$.

To show that every eigenvalue of A_r is an eigenvalue of $\tilde{A}_r$, suppose λ is a non-zero eigenvalue of A_r with eigenvector φ. Let $w = C^\dagger\varphi$. Then we have

$$\begin{aligned}
\tilde{A}_r w &= (C^\dagger B)(C^\dagger\varphi) \\
&= C^\dagger A_r\varphi \\
&= \lambda C^\dagger\varphi \\
&= \lambda w.
\end{aligned}$$

Moreover, $w \neq 0$, because if $w = C^\dagger\varphi = 0$, then $BC^\dagger\varphi = A_r\varphi = \lambda\varphi = 0$. So $\lambda = 0$, which is a contradiction. So λ is an eigenvalue of $\tilde{A}_r$ with eigenvector w. □

Let P_{X_r} be the orthogonal projection operator onto the column space of X_r. Note that the column space of X_r is equal to the column space of C because C consists of r linearly independent rows of X_r. So

$$P_{X_r} = P_C = CC^\dagger.$$

Theorem 101. *Let $\hat{\varphi}$ be defined as in Equation* (3.7)*, and let φ be defined as in Equation* (3.8)*. Then $\hat{\varphi}$ is an eigenvector of $P_{X_r}A_r$ with eigenvalue λ. Moreover, $\hat{\varphi} = P_{X_r}(\varphi)$.*

Proof. To show $\hat{\varphi}$ is an eigenvector of $P_{X_r}A_r$, we have

$$\begin{aligned}
P_{X_r}(A_r\hat{\varphi}) &= (CC^\dagger)(BC^\dagger)(Cw) \\
&= C(C^\dagger B)I_r w \\
&= C\tilde{A}_r w \\
&= \lambda C w \\
&= \lambda\hat{\varphi}.
\end{aligned}$$

So λ is an eigenvalue of $\hat{\varphi}$. Moreover, we have

$$\begin{aligned}
P_{X_r}(\varphi) &= (CC^\dagger)(\frac{1}{\lambda}Bw) \\
&= \frac{1}{\lambda}C\tilde{A}_r w \\
&= Cw \\
&= \hat{\varphi}
\end{aligned}$$

□

Again, we assume A_r will be a suitable substitute for A for calculating the dynamics of the system. We can find $\tilde{z}_{n+N}$, an approximation of the data vector z_{n+N} for $N > 0$ by calculating

$$\begin{aligned}
\tilde{z}_{n+N} &= (P_{X_r}A_r)^N P_{X_r} z_n \\
&= (CC^\dagger A_r)^N (CC^\dagger) z_n \\
&= C\tilde{A}_r^N C^\dagger z_n.
\end{aligned}$$

We use the fact that

$$\begin{aligned} P_{X_r} &= X_r X_r^\dagger \\ &= C\hat{X}R^\dagger R\hat{X}^{-1}C^\dagger \\ &= CC^\dagger. \end{aligned}$$

Note that Theorem 100 and Theorem 101 hold for any rank r approximation $X_r = LS$, where L is full rank $m \times r$ and S is full rank $r \times n$. When X_r is a low-rank approximation obtained by the SVD, then $L = U$ and $S = \Sigma V^*$. When X_r is a low-rank approximation obtained by the skeleton decomposition, then $L = C$ and $S = \hat{X}^{-1}R$.

No assumptions on the error of the low-rank approximation X_r are made. In practice, X_r should be close to X so A_r closely approximates the dynamics generated by A. When a low-rank approximation obtained by the skeleton decomposition with respect to the submatrix $\hat{X}$, a maximum volume algorithm such as Algorithm 10 or an alternating version of Algorithm 12 should be used to choose $\hat{X}$ so the error of the low-rank approximation is minimized by Theorem 69.

Example 102. *We compare dynamic mode decompositions using low-rank approximation obtained by the SVD and the maximum volume skeleton decomposition (MVSD) by testing on one dimensional low-noise temperature data in time. The matrix of data* $Z = [z_0 \ \cdots \ z_{2000}]$ *is* 38×2001*, meaning there are* 38 *spacial degrees of freedom and* 2001 *time steps. Let* $X = [z_0 \ \cdots \ z_{1999}]$ *be the* 38×2000 *submatrix consisting of all but the last column and let* $Y = [z_1 \ \cdots \ z_{2000}]$ *be the* 38×2000 *submatrix consisting of all but the first column. Let* X_{10} *be a rank ten approximation of* X *using the MVSD or SVD. Define* $A_{10} = YX_{10}^\dagger$*. When* X_{10} *is obtained by the SVD as in Equation* (3.1),

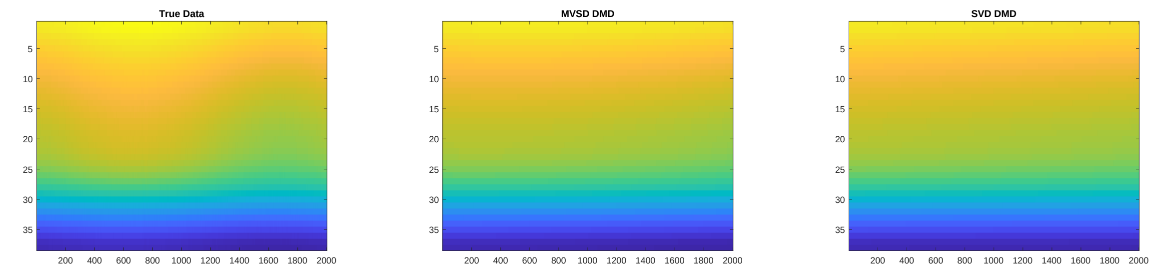
$$\tilde{z}_k = U_{10}\tilde{A}_{10}^k U_{10}^* z_0.$$

The vectors $\tilde{z}_k$ *are assembled into the* 38×2001 *matrix* $\tilde{Z}_{SVD}$ *such that the kth column of* $\tilde{Z}_{SVD}$ *is* $\tilde{z}_k$*.*

When X_{10} is defined with the MVSD as in Equation (3.4),

$$\tilde{z}_k = C\tilde{A}_{10}^k C^\dagger z_0.$$

The vectors $\tilde{z}_k$ are assembled into the 38×2001 *matrix $\tilde{Z}_{MVSD}$ such that the kth column of $\tilde{Z}_{MVSD}$ is $\tilde{z}_k$.*



(a) Z, low-noise data matrix (b) $\tilde{Z}_{\text{MVSD}}$, relative error 0.0298 (c) $\tilde{Z}_{\text{SVD}}$, relative error 0.0350

Figure 3.2: Data approximated by the DMD with $r = 10$.

As can be seen in Figure 3.2, the dynamic mode decomposition using either the SVD or MVSD can accurately predict the dynamics of the data.

3.3 Simulation Data Compression

It is estimated that ITER facility operations will produce upwards of a petabyte continuous stream of diagnostic information per day resulting in exascale data handling requirements annually (Lister et al., 2003). Application of mathematical techniques that can facilitate large-scale data compression are required for efficient analysis of burning plasma research.

We present numerical results on applying the skeleton decomposition as defined in Definition 68 with respect to a submatrix chosen by an alternating maximum volume skeleton decomposition presented in Algorithm 10, and an alternating version of the greedy maximum volume skeleton decomposition presented in Algorithm 12, to the plasma simulation data produced by SOLPS-ITER for the experimental DIII-D device (Petty et al., 2019).

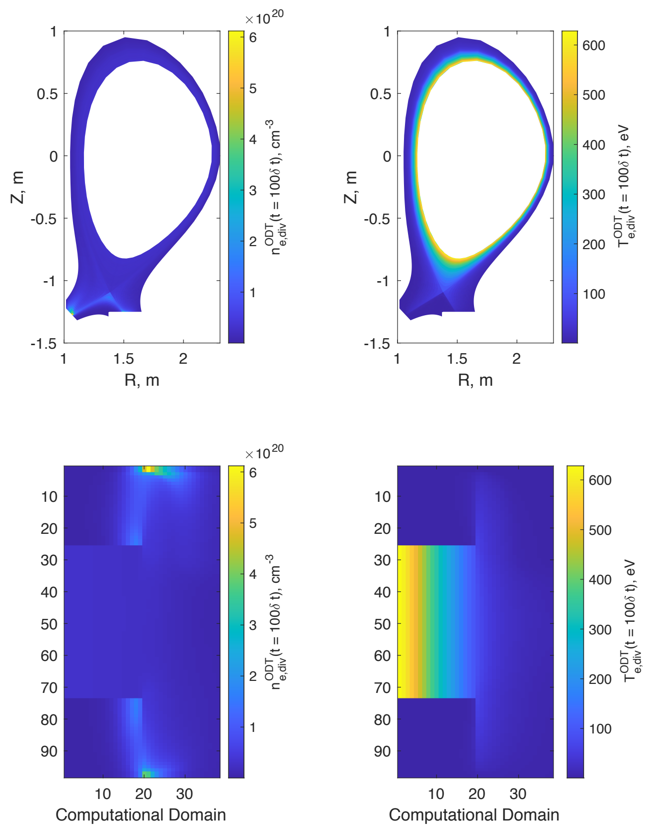
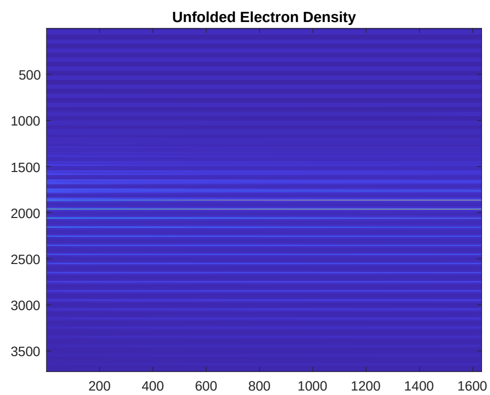
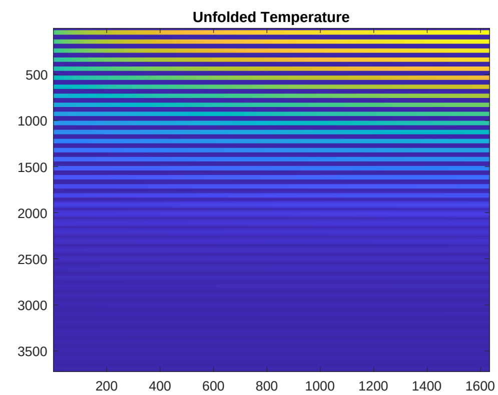


Figure 3.3: A two dimensional plasma simulation of electron density (left) and temperature (right) at a fixed time. The entries in the matrix are mapped to the corresponding physical location in the device.

The simulation data consists of electron density and plasma temperature over two spacial dimensions and time. See Figure 3.3 for an example spacial slice. There are 98×38 spacial degrees of freedom, and 1633 time steps giving a $98 \times 38 \times 1633$ array of data. Each 98×38 spacial slice is vectorized resulting in 3724×1633 matrices, see Figure 3.4.



(a) Unfolded electron density matrix

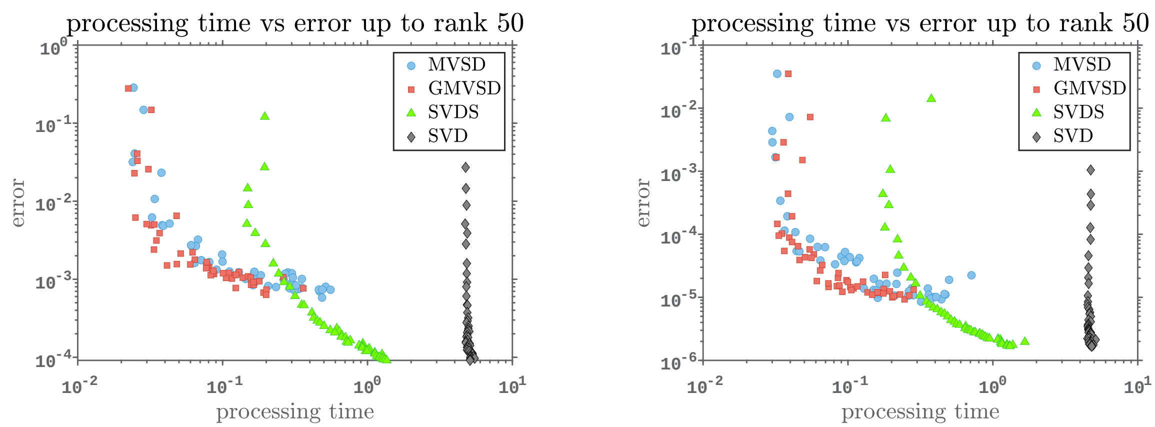


(b) Unfolded temperature matrix

Figure 3.4: 3724×1633 simulation data unfolded into matrices. Each column is a vectorized 98×38 spacial slice.

We compare the error and the processing time for finding a rank r approximation of simulation data shown in Figure 3.4 for $1 \leq r \leq 50$ in Figure 3.5. To calculate the rank r approximation, we use the MVSD, the greedy maximum volume skeleton decomposition (GMVSD), the SVDS which computes the first r singular values and left and right singular vectors, and the SVD. The further left a node is, the shorter the processing time. The further down a node is, the lower the error. The processing time to compute the SVD remains constant because every singular vector and singular value is being computed regardless of the rank. For a fixed rank r, the error of a rank r approximation will be smaller when using the SVD compared to the MVSD because the SVD calculates a closest rank r approximation.

If there is a large enough acceptable error tolerance, the MVSD and GMVSD are much faster to compute than the SVD and the SVDS for the same error. For example, if a relative error of 10^{-4} is acceptable to compress the temperature data, then the MVSD or GMVSD can compute a low-rank approximation with a relative error of 10^{-4} faster than the SVD or SVDS.



(a) Electron density

(b) Temperature

Figure 3.5: Each node corresponds to a rank from 1 to 50 and relative error is used. The data being compressed is two dimensional electron density and temperature in time. If a large enough error is acceptable for compression, then the MVSD outperforms the SVD and SVDS.

Chapter 4

Low-Rank Tensor Completion

The data completion methods we have discussed so far have been applicable to incomplete matrices, or two-dimensional arrays. However, data is often parameterized by more than two dimensions. For example, in section Section 3.3, simulated data is given in a three dimensional array which consisted of two spacial dimensions and one time dimension.

Definition 103. *Arrays of arbitrary dimension are referred to as* tensors*. The dimension of the array is referred to as the* order *of the tensor. For example, matrices are order two tensors. Let U be a vector space of dimension m with basis $u_1, \dots, u_m$ and let V be a vector space of dimension n with basis $v_1, \dots, v_n$. The tensor product $U \otimes V$ is the vector space of dimension mn spanned by tensors of the form $u_i \otimes v_j$ with the following relations. For elements u, u', v, v', an scalar λ, we have*

1. $\lambda(u \otimes v) = (\lambda u) \otimes v = u \otimes (\lambda v)$
2. $(u + u') \otimes v = u \otimes v + u' \otimes v$
3. $u \otimes (v + v') = u \otimes v + u \otimes v'$.

If W is a vector space of dimension p with basis $w_1, \dots, w_p$, then $U \otimes V \otimes W$ is a vector space of dimension mnp spanned by tensors of the form $u_i \otimes v_j \otimes w_k$. A tensor $T \in U \otimes V \otimes W$ may be expressed as a sum

$$T = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}(u_i \otimes v_j \otimes w_k)$$

with respect to this basis. T may also be expressed as the third order array of coefficients

$$T = [a_{ijk}]$$

where $1 \leq i \leq m$, $1 \leq j \leq n$, and $1 \leq k \leq p$. The Euclidean norm of T is defined as

$$\|T\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}^2.$$

Given a tensor product of vector spaces $V_1 \otimes \cdots \otimes V_d$, an element $T \in V_1 \otimes \cdots \otimes V_d$ is a tensor of order d. In this chapter we express definitions and theorems in terms of order three tensors for notational convenience. However, many definitions and theorems may be generalized to higher order tensors.

4.1 Mathematical Preliminaries

To generalize methods of low-rank matrix completion to methods for tensor completion, we must first generalizing the notion of the rank of a matrix. There are various ways to define the rank of a tensor. In this section, we present some definitions and known results on the geometry of low-rank tensors.

Let U, V be vector spaces. Let U^* denote the dual vector space of U. Let $u^* \in U^*$, and $v \in V$. Then $u^* \otimes v \in U^* \otimes V$, and one can define the rank one linear map

$$u^* \otimes v : U \to V$$

$$(u^* \otimes v)(x) = u^*(x)v.$$

In general, for $T \in U^* \otimes V$, the rank of the linear map $T : U \to V$ is the smallest r such that there exists $u_1^*, \dots, u_r^* \in U^*$ and $v_1, \dots, v_r \in V$ such that $T = \sum_{s=1}^{r} u_s^* \otimes v_s$. Defining the rank this way generalizes to higher order tensors.

Definition 104. *An element $T \in U \otimes V \otimes W$ is called* rank one *if $T = u \otimes v \otimes w$, for some $u \in U$, $v \in V$, and $w \in W$. The* rank *of a tensor T is the smallest r such that T is the sum of r rank one tensors, and is denoted* $\operatorname{rank}(T)$.

While this is a straightforward way to generalize the definition of rank, many nice properties do not generalize. It is useful to introduce alternative generalizations of rank. Because a tensor $T \in U^* \otimes V$ represents a linear map $T : U \to V$ and the double dual U^{**} is canonically isomorphic to U, a tensor $T \in U \otimes V$ represents a linear map $T : U^* \to V$. For a tensor $T = \sum_s u_s \otimes v_s$ Define $T(u^*) = \sum_s u^*(u_s)v_s$.

More generally, a tensor $T \in U \otimes V \otimes W$ may be unfolded as linear maps

$$\begin{aligned} T^{(1)} &:(V \otimes W)^* \to U \\ T^{(2)} &:(U \otimes W)^* \to V \\ T^{(3)} &:(U \otimes V)^* \to W. \end{aligned}$$

where if $T = \sum_s u_s \otimes v_s \otimes w_s$, then

$$\begin{aligned} T^{(1)}(f) &= \sum_s f(v_s \otimes w_s)u_s \\ T^{(2)}(f) &= \sum_s f(u_s \otimes w_s)v_s \\ T^{(3)}(f) &= \sum_s f(u_s \otimes v_s)w_s \end{aligned}$$

The linear map $T^{(i)}$ is called the *mode-i unfolding* or *flattening* of T.

Definition 105. *The* multilinear rank*, also known as the* Tucker rank*, of* $T \in U \otimes V \otimes W$ *is denoted* $\mu\text{rank}(T)$*, and is defined to be the* 3*-tuple of natural numbers*

$$\mu\text{rank}(T) = \big(\text{rank}(T^{(1)}), \text{rank}(T^{(2)}), \text{rank}(T^{(3)})\big)$$

Note that for an order two tensor $T \in U \otimes V$, $\text{rank}(T^{(1)}) = \text{rank}(T^{(2)})$ because the row rank of a matrix is equal to its column rank.

Proposition 106. *If* T *is an* $m \times n \times p$ *tensor such that* $\mu\text{rank}(T) = (r_1, r_2, r_3)$*, then*

$$r_1 \leq \min(m, np)$$
$$r_2 \leq \min(n, mp)$$
$$r_3 \leq \min(p, mn).$$

This is because $T^{(1)}$ *is an* $m \times np$ *matrix,* $T^{(2)}$ *is an* $n \times mp$ *matrix, and* $T^{(3)}$ *is an* $p \times mn$ *matrix.*

Given basis $\{u_i \otimes v_j \otimes w_k\}$ for $U \otimes V \otimes W$, we may express T as a linear combination of basis elements

$$T = \sum_{k=1}^{p} \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ijk}(u_i \otimes v_j \otimes w_k).$$

We may represent T as the order three array of coefficients $[a_{ijk}]$. Then $T^{(i)}$ may be expressed as a matrix by unfolding the slices of $[a_{ijk}]$ along the ith coordinate.

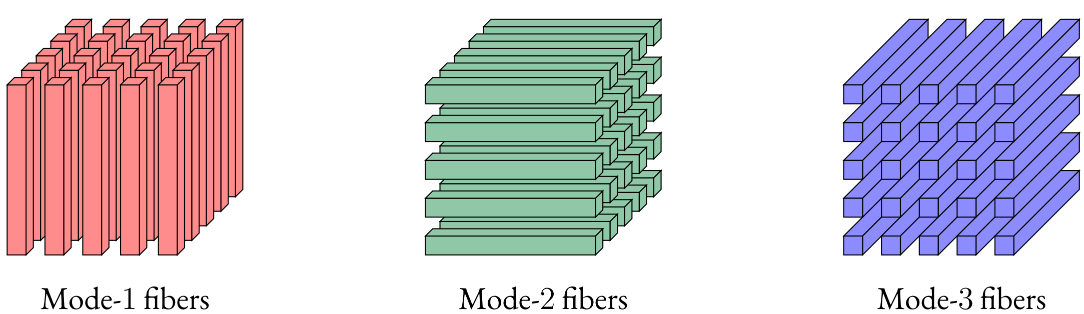


Figure 4.1: Fibers of an order three tensor

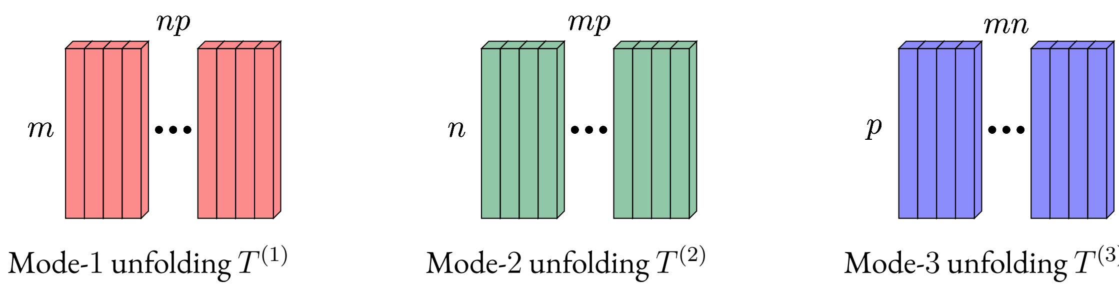


Figure 4.2: Unfoldings of an $m \times n \times p$ tensor T

The first, second, and third components of the multilinear are equal to the maximum number of linearly independent mode-1 (column) fibers, mode-2 (row) fibers, and mode-3 (tube) fibers respectively.

The rank and the multilinear rank of tensors are related in the following way.

Lemma 107. *(Landsberg, 2011) For any tensor* T*,* $\max(\mu\text{rank}(T)) \leq \text{rank}(T)$.

Proof. Given a tensor $T \in U \otimes V \otimes W$, suppose $\text{rank}(T) = r$. Then T may be expressed as the sum of r rank one tensors $T = \sum_{s=1}^{r} Z_s$ where Z_s is a rank one tensor for all s. Because Z_s is rank one,

$$Z_s = u_s \otimes v_s \otimes w_s$$

for some u_s, v_s, and w_s. So

$$Z_s^{(1)}(f) = f(v_s \otimes w_s)u_s$$

where $f \in (V \otimes W)^*$. The image of $Z_s^{(1)}$ is contained in the span of $u_s \in U$, so

$$\text{rank}(Z_s^{(1)}) \leq 1$$

for all s, which implies

$$\begin{aligned}\operatorname{rank}(T^{(1)}) &= \operatorname{rank}(\sum_{s=1}^{r} Z_s^{(1)}) \\ &\leq \sum_{s=1}^{r} \operatorname{rank}(Z_s^{(1)}) \\ &\leq r,\end{aligned}$$

and similarly for $T^{(2)}$ and $T^{(3)}$. So each component of the multilinear rank is at most r. □

One issue with the definition for the rank of a tensor is that, unlike the set of matrices with rank at most r, the set of tensors with rank at most r may not be closed. In other words, it may be possible to express a high rank tensor as the limit of low-rank tensors. To study the closure of the set of tensors with rank at most r, we introduce the notion of border rank.

Definition 108. *The* border rank *of a tensor T, denoted $\overline{\operatorname{rank}}(T)$, is the smallest r such that there exists a sequence of rank r tensors $\{T_\lambda\}$ such that $\lim_{\lambda\to\infty} T_\lambda = T$. Equivalently, the border rank of a tensor T is the smallest r such that there exists a tensor of rank r in the ball $B_\epsilon(T) = \{X \mid \|X - T\| < \epsilon\}$ for all $\epsilon > 0$.*

Note that the border rank of a tensor is no more than the rank of a tensor. That is,

$$\overline{\operatorname{rank}}(T) \leq \operatorname{rank}(T).$$

We give an example of a tensor whose rank and border rank differ.

Theorem 109. *(Landsberg, 2011) Let $T = u \otimes u \otimes v + u \otimes v \otimes u + v \otimes u \otimes u$, where u and v are linearly independent. Then $\operatorname{rank}(T) = 3$. However, $\overline{\operatorname{rank}}(T) = 2$. To see this, let*

$$T_\lambda = \lambda(u + \frac{1}{\lambda}v) \otimes (u + \frac{1}{\lambda}v) \otimes (u + \frac{1}{\lambda}v) - \lambda(u \otimes u \otimes u).$$

Then $\operatorname{rank}(T_\lambda) = 2$, *and* $\lim_{\lambda\to\infty} T_\lambda = T$.

We will show $\operatorname{rank}(T) = 3$ by contradiction in the case where $u = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $v = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$.

Proof. Let

$$T = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}.$$

Suppose $\operatorname{rank}(T) \leq 2$. Then we may express

$$T = \begin{bmatrix}u_{11}\\u_{12}\end{bmatrix} \otimes \begin{bmatrix}v_{11}\\v_{12}\end{bmatrix} \otimes \begin{bmatrix}w_{11}\\w_{12}\end{bmatrix} + \begin{bmatrix}u_{21}\\u_{22}\end{bmatrix} \otimes \begin{bmatrix}v_{21}\\v_{22}\end{bmatrix} \otimes \begin{bmatrix}w_{21}\\w_{22}\end{bmatrix}.$$

for some $u_{11}, u_{12}, u_{21}, u_{22}, v_{11}, v_{12}, v_{21}, v_{22}, w_{11}, w_{12}, w_{21}$, and w_{22}. By expanding each expression into a $2 \times 2 \times 2$ array and equating them, we get the system of eight equations

$$\begin{aligned}
u_{11}v_{11}w_{11} + u_{21}v_{21}w_{21} &= 0 & \qquad u_{12}v_{11}w_{11} + u_{22}v_{21}w_{21} &= 1\\
u_{11}v_{11}w_{12} + u_{21}v_{21}w_{22} &= 1 & \qquad u_{12}v_{11}w_{12} + u_{22}v_{21}w_{22} &= 0\\
u_{11}v_{12}w_{11} + u_{21}v_{22}w_{21} &= 1 & \qquad u_{12}v_{12}w_{11} + u_{22}v_{22}w_{21} &= 0\\
u_{11}v_{12}w_{12} + u_{21}v_{22}w_{22} &= 0 & \qquad u_{12}v_{12}w_{12} + u_{22}v_{22}w_{22} &= 0.
\end{aligned}$$

One may use computer algebra software to verify that this system of equations has no solutions. □

An open question is, given a tensor T of border rank r, what is the largest rank the tensor can have? For example, it is known that a border rank four tensor in $\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4$ can have rank at most seven, but in general the question is open (Landsberg & Michałek, 2017).

Let σ_r denote the set of tensors in $U \otimes V \otimes W$ with rank at most r.

Definition 110. *Let the* $B_\epsilon(T) = \{X \in U \otimes V \otimes W \mid \|X - T\| < \epsilon\}$ *denote the ball centered at* T *with radius* ϵ*. The* generic rank *of a tensor* T*, denoted* $\mathrm{grank}(T)$*, is the least* r *such that the intersection* $B_\epsilon(T) \cap \sigma_r$ *is of positive measure for all* $\epsilon > 0$*.*

Note that if r is the generic rank of T, there exists a sequence $\{T_\lambda\}$ where $\mathrm{rank}(T_\lambda) = r$ such that $\lim_{\lambda\to\infty} T_\lambda = T$. Therefore, the border rank of T is no larger than the generic rank of T. That is, $\overline{\mathrm{rank}}(T) \leq \mathrm{grank}(T)$. In general, the rank of a tensor T could be larger or smaller than the generic rank of T. Over $\mathbb{C}$, each tensor in $U \otimes V \otimes W$ has the same generic rank r (Landsberg, 2011). Moreover, r is equal to the maximum border rank over all tensors in $U \otimes V \otimes W$ (Strassen, 1983). In this case r is referred to as the generic, or typical, rank of $U \otimes V \otimes W$. However, over $\mathbb{R}$ the generic rank may not be unique. A rank r is called a *typical rank* if the set of tensors of rank r in $U \otimes V \otimes W$ has non-zero measure. Equivalently, r is a typical rank if it is the generic rank of some tensor T. Again, over $\mathbb{C}$ there is a unique typical rank, but over $\mathbb{R}$ there may be more than one typical rank.

In $M_{m\times n}$, the space of $m \times n$ matrices, the only typical rank is $\min(m, n)$ over $\mathbb{R}$ or $\mathbb{C}$, which is also the maximum rank any m by n matrix can have. In general over $\mathbb{C}$, there could be ranks larger than a typical rank such that the set of tensors with rank r is non-empty and has measure zero. It is known that the typical rank in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ is two. In $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$, both two and three are typical ranks. More generally, in $\mathbb{R}^m \otimes \mathbb{R}^n \otimes \mathbb{R}^p$, every rank from the largest border rank to the largest rank is a typical rank (Friedland, 2012).

Given vector spaces U, V, and W, let $m = \dim(U)$, $n = \dim(V)$, and $p = \dim(W)$. Note that the set of rank one tensors in $U \otimes V \otimes W$ has dimension $m + n + p - 2$ because it is parameterized by elements $u \in U, v \in V, w \in W$ up to scale plus one scalar. Rank r tensors are sums of r rank one tensors, so as shown in (Strassen, 1983), the space of rank at most r tensors has bounded dimension

$$\dim(\sigma_r) \leq r(m + n + p - 2). \tag{4.1}$$

To approximate the generic rank over $\mathbb{C}$, note that $\dim(U \otimes V \otimes W) = mnp$. By substituting mnp into the left hand side of Inequality (4.1) we conclude that if r is the typical rank,

$$\left\lceil \frac{mnp}{m+n+p-2} \right\rceil \leq r.$$

Define the *expected rank* of a tensor in $\mathbb{C}^m \otimes \mathbb{C}^n \otimes \mathbb{C}^p$ to be $\left\lceil \frac{mnp}{m+n+p-2} \right\rceil$.

In general, there is no easy way to calculate the generic rank explicitly. However, it is known in some cases. For example, it is known that for all $n \neq 3$, the generic rank of $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ is the expected rank $\left\lceil \frac{n^3}{3n-2} \right\rceil$. When $n = 3$, the generic rank is five (Lickteig, 1985).

It is useful to know the typical ranks because an incomplete tensor with entries chosen from a continuous density will have a rank r completion with non-zero probability only if r is a typical rank.

Because there are several definitions for the rank of a tensor, there are several ways to define spaces of rank at most r tensors. Recall that σ_r denotes the space of tensors in $U \otimes V \otimes W$ of rank at most r. That is,

$$\sigma_r = \{X \in U \otimes V \otimes W \mid \operatorname{rank}(X) \leq r\}.$$

Let $\hat{\sigma}_r$ denote the set of tensors of border rank at most r in $U \otimes V \otimes W$. That is,

$$\hat{\sigma}_r = \{X \in U \otimes V \otimes W \mid \overline{\operatorname{rank}}(X) \leq r\}.$$

By Definition 108, $\hat{\sigma}_r$ is the closure of σ_r. In fact, $\hat{\sigma}_r$ is also the Zariski closure of σ_r. In other words, $\hat{\sigma}_r$ is an algebraic variety, meaning it is the zero set of a system of polynomials equations. In general, the polynomials that generate $\hat{\sigma}_r$ are not completely known. We recall some known results. Other known results can be found in (Landsberg, 2011).

Definition 111. *The* subspace variety $\hat{S}ub_r \subset U \otimes V \otimes W$ *is the space of tensors such that each entry of the multilinear rank is at most* r. *In other words, a tensor* X *is in* $\hat{S}ub_r$ *if and only if* $\operatorname{rank}(X^{(i)}) \leq r$ *for* *all* i*. So we have*

$$\begin{aligned}\hat{S}ub_r &= \{X \in U \otimes V \otimes W \mid \max(\mu\text{rank}(X)) \leq r\} \\ &= \{X \in U \otimes V \otimes W \mid \text{rank}(X^{(i)}) \leq r, i = 1, 2, 3\}.\end{aligned}$$

Recall that a matrix M has rank at most r if and only if all $(r+1) \times (r+1)$ minors of M vanish. So a tensor X is in $\hat{S}ub_r$ if and only if $\text{rank}(X^{(i)}) \leq r$ for all i, if and only if all $(r+1) \times (r+1)$ minors of $X^{(i)}$ vanish for each i. Therefore, $\hat{S}ub_r$ is an algebraic variety generated by the $(r+1) \times (r+1)$ minors of the mode-1, mode-2, and mode-3 unfoldings. These minors set equal to zero are called *equations of flattening*.

Theorem 112. *(Landsberg, 2011) The spaces* σ_r*,* $\hat{\sigma}_r$*, and* $\hat{S}ub_r$ *are related in the sense that*

$$\sigma_r \subset \hat{\sigma}_r \subset \hat{S}ub_r$$

in other words,

$$\max(\mu\text{rank}(T)) \leq \overline{\text{rank}}(T) \leq \text{rank}(T)$$

for all $T \in U \otimes V \otimes W$*.*

Proof. Recall from Lemma 107 that if the rank of a tensor T is at most r, then the maximum component of the multilinear rank of T is at most r. That is, $\text{rank}(T) \leq r$ implies $\max(\mu\text{rank}(T)) \leq r$. Therefore if $\text{rank}(T) \leq r$, then the $(r+1) \times (r+1)$ minors of every mode-i unfolding of T vanish. In other words, we have $\sigma_r \subset \hat{S}ub_r$. Because $\hat{S}ub_r$ is an algebraic variety, it is a closed set, so the closure of σ_r, that is $\hat{\sigma}_r$, is also a subset of $\hat{S}ub_r$. So we have

$$\sigma_r \subset \hat{\sigma}_r \subset \hat{S}ub_r.$$

□

We noted in the proof of Theorem 112 that if the rank of a polynomial is at most r, the $(r+1)\times(r+1)$ minors of the mode-i unfolding of a tensor T vanish. Because polynomials are continuous, if a sequence of tensors each satisfy a polynomial equation, the limit must also satisfy that polynomial equation. It follows that the $(r+1)\times(r+1)$ minors of the mode-i unfolding of a tensor T with border rank at most r must vanish.

$T \in \hat{S}ub_r$ is a necessary, but not always a sufficient condition to imply $T \in \hat{\sigma}_r$. However, it is sufficient when $r = 1$. Moreover, it is sufficient to imply $T \in \sigma_1$. More specifically, we have the following.

Proposition 113. *(Landsberg, 2011) The set of rank at most one, border rank at most one, and maximum multilinear rank at most one tensors are equal.*

$$\sigma_1 = \hat{\sigma}_1 = \hat{S}ub_1.$$

In particular, this implies that σ_r is closed when $r = 1$.

We now discuss how to explicitly compute the various ranks of a tensor. The easiest type of rank to compute is the multilinear rank. Simply compute the rank of the mode-i unfolding matrices $T^{(i)}$ for all i.

The border rank is difficult to compute in general. The set of tensors of border rank at most r is an algebraic variety, so there exists a finite number of polynomials such that if T is a zero of all such polynomials, then T has border rank at most r. In theory, we could test if T is a zero of each polynomial equation. If it is, then T has border rank at most r. However, the polynomials that generate $\hat{\sigma}_r$ are not all known in general.

To compute the rank of a tensor $T \in U \otimes V \otimes W$, we need to find the smallest r such that T can be written as a sum of r rank one tensors. To check if T can be written as a sum of r rank one tensors, assume we can express $T = \sum_{s=1}^{r} u_s \otimes v_s \otimes w_s$ for some u_s, v_s, and w_s. Express T as a third order array where

$$T_{ijk} = \sum_{s=1}^{r} (u_s)_i (v_s)_j (w_s)_k.$$

This will give mnp equations, which may be solved using Gröbner bases for example. If there is a solution, then T has rank at most r. If there is no solution then T has rank greater than r.

Given an $m \times n \times p$ tensor T, if T can be approximated well by a rank r tensor for some r, it is much more efficient to store as sum of r rank one tensors. In particular, storing T requires storing mnp entries. However, storing r rank one tensors only requires storing $r(m+n+p)$ entries. It is therefore often desirable to calculate a closest rank r approximation of T. That is, we would like to find a solution to the minimization problem

$$\min_{u_i,v_i,w_i} \left\| T - \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i \right\|$$

$$\textit{s.t. } u_i \in U, v_i \in V, w_i \in W$$

Recall from Theorem 11 that the problem of finding a closest rank r approximation is solved for matrices by zeroing out small singular values. However, in the case of higher order tensors, the problem is ill-posed (De Silva & Lim, 2008). This is because σ_r, the set of tensors of rank at most r, is not a closed set. Recall from Theorem 109 that T is rank three, and there is a sequence of rank two tensors T_λ such that $\lim_{\lambda\to\infty} T_\lambda = T$. So, T does not have a closest rank two approximation because there are rank two tensors arbitrarily close to T.

Moreover, the norms of the individual rank one terms are unbounded. That is,

$$\lim_{\lambda\to\infty} \left\| \lambda(u + \frac{1}{\lambda}v) \otimes (u + \frac{1}{\lambda}v) \otimes (u + \frac{1}{\lambda}v) \right\| = \infty$$

$$\lim_{\lambda\to\infty} \|\lambda(u \otimes u \otimes u)\| = \infty.$$

This phenomenon is sometimes called the problem of diverging components (Stegeman, 2008). One approach to avoid this problem when finding a low-rank approximation is to impose additional constraints that bound the norms of the rank one components.

4.2 Exact Tensor Completion

Like matrix completion, the problem of tensor completion is, given a partially known tensor T_Ω, complete the unknown entries of T_Ω subject to the constraint that the resulting tensor is low-rank. There are multiple notions of what low-rank could mean for tensors. For example the tensor could have low-rank, low-border rank, or low-multilinear-rank.

Let U, V, and W be vector spaces of dimension $m, n,$ and p respectively. Let T_Ω denote a partially known tensor. Let $\Omega \subset [m] \times [n] \times [p]$ denote the set of positions of the known entries of T_Ω. Let $P_\Omega : U \otimes V \otimes W \to U \otimes V \otimes W$ denote the orthogonal projection map where $P_\Omega(X)$ sets entries with positions not in Ω equal to zero. Also let $\mathcal{A}_\Omega = P_\Omega^{-1}(T_\Omega)$ denote the linear variety of completions of T_Ω. The objective of low-rank tensor completion is to find a solution to a minimization problem

$$\begin{aligned} &\min_{T \in U\otimes V\otimes W} \operatorname{rank}(T) && \min_{T \in U\otimes V\otimes W} \overline{\operatorname{rank}}(T) && \min_{T \in U\otimes V\otimes W} \max(\mu\operatorname{rank}(T)) \\ &\textit{s.t. } P_\Omega(T) = T_\Omega && \textit{s.t. } P_\Omega(T) = T_\Omega && \textit{s.t. } P_\Omega(T) = T_\Omega. \end{aligned}$$

The set of tensors of rank at most r is not closed and the rank of a tensor is difficult to compute, so there are issues with minimizing the rank of a tensor. The issue with minimizing the border rank is that the border rank is difficult to compute.

In this section, we focus on completing partially known tensors with respect to the constraint that each component of the multilinear rank is at most r. Recall from Theorem 112 that there is a relationship between rank, border rank, and multilinear rank. In particular,

$$\sigma_r \subset \hat{\sigma}_r \subset \hat{S}ub_r.$$

Therefore, given a partially known tensor T_Ω, if there is a unique maximum border rank at most r completion of T_Ω, then either there is a unique rank at most r or border rank at most r completion, or none exist. Moreover, if they exist, then they are equal to the completion with respect to the multilinear rank.

More formally, we have the following theorem.

Theorem 114. *Given a partially known tensor T_Ω, suppose $\mathcal{A}_\Omega \cap \sigma_r$ is non-empty. That is, suppose there exists at least one rank at most r completion of T_Ω. Suppose there exists a unique multilinear rank at most r completion T of T_Ω. That is, suppose $\mathcal{A}_\Omega \cap \hat{S}ub_r = \{T\}$ for some T. Then*

$$\mathcal{A}_\Omega \cap \sigma_r = \mathcal{A}_\Omega \cap \hat{\sigma}_r = \{T\}.$$

Proof. By Theorem 112, $\sigma_r \subset \hat{\sigma}_r \subset \hat{S}ub_r$. Intersecting with $\mathcal{A}_\Omega$, we have

$$\mathcal{A}_\Omega \cap \sigma_r \subset \mathcal{A}_\Omega \cap \hat{\sigma}_r \subset \mathcal{A}_\Omega \cap \hat{S}ub_r = \{T\}.$$

Then by assumption, because there is at least one rank or border rank at most r completion, we must have $\mathcal{A}_\Omega \cap \sigma_r = \mathcal{A}_\Omega \cap \hat{\sigma}_r = \{T\}$. □

We introduce sufficient conditions for an incomplete order three tensor T_Ω to have a unique multilinear rank (r, r, r) completion under the assumption that at least one exists. As a special case, we introduce sufficient conditions for an incomplete tensor to have a unique multilinear rank $(1, 1, 1)$ completion.

Theorem 115. *Given T_Ω, suppose $T_{i,1,1}$ is known for all i, $T_{1,j,1}$ is known for all j, and $T_{1,1,k}$ is known for all k, and suppose $T_{1,1,1} \neq 0$. Then if $\mathcal{A}_\Omega \cap \hat{S}ub_1$ is non-empty, T_Ω has a unique rank one completion.*

Proof. Because $\hat{S}ub_1 = \{T \in U \otimes V \otimes W \mid \text{rank}(T^{(i)}) \leq 1, i = 1, 2, 3\}$, the polynomials that generate $\hat{S}ub_1$ are all 2×2 minors of the mode-1, mode-2, and mode-3 unfoldings. In other words, a tensor T is in $\hat{S}ub_1$ if and only if all 2×2 minors vanish in each of the unfoldings

$$T^{(1)} : (V \otimes W)^* \to U$$

$$T^{(2)} : (U \otimes W)^* \to V$$

$$T^{(3)} : (U \otimes V)^* \to W.$$

By assumption, the entry T_{ijk} in T_Ω is known if exactly two or more indices in (i, j, k) are equal to one. We first show that we can recover all entries with one index in (i, j, k) equal to one. That is, we can recover all entries of the form T_{1jk}, T_{i1k}, or T_{ij1}.

The equations of flattening give us equations of the form

$$\begin{vmatrix} T_{111} & T_{11k} \\ T_{1j1} & T_{1jk} \end{vmatrix} = 0$$

$$\begin{vmatrix} T_{111} & T_{11k} \\ T_{i11} & T_{i1k} \end{vmatrix} = 0$$

$$\begin{vmatrix} T_{111} & T_{1j1} \\ T_{i11} & T_{ij1} \end{vmatrix} = 0$$

for all i, j, and k. Here the bottom-right entry of each matrix is unknown and all other entries are known. Moreover, because $T_{111} \neq 0$, we may solve for each unknown entry which is equal to

$$T_{1jk} = T_{11k}(T_{111}^{-1})T_{1j1}$$

$$T_{i1k} = T_{11k}(T_{111}^{-1})T_{i11}$$

$$T_{ij1} = T_{i11}(T_{111}^{-1})T_{1j1}.$$

Next, we may complete an arbitrary entry T_{ijk} by considering equations of flattening of the form

$$\begin{vmatrix} T_{111} & T_{1jk} \\ T_{i11} & T_{ijk} \end{vmatrix} = 0$$

and solving $T_{ijk} = T_{1jk}(T_{111}^{-1})T_{i11}$.

Because every maximum multilinear rank at most r completion of T_Ω must satisfy the equations used to complete T_Ω, and each equation had a unique solution, the constructed completion is the unique maximum multilinear rank at most r completion of T_Ω. Moreover, by Proposition 113, $\sigma_1 = \hat{\sigma}_1 = \hat{S}ub_1$. So T_Ω also has a unique rank one completion which is equal to the constructed completion. $\square$

We give an example of a partially known tensor which can be uniquely completed by Theorem 115.

Example 116. *Consider Ω defined as in Theorem 115, and suppose $T_{ijk} = 1$ for all $(i, j, k) \in \Omega$. Then T_Ω has the unique rank one, border rank one, and multilinear rank $(1, 1, 1)$ completion T, where $T_{ijk} = 1$ for all (i, j, k). T may also be written in the rank one form*

$$T = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

We now give the more general version of Theorem 115 for $r \geq 1$.

Theorem 117. *Given a partially known $m \times n \times p$ tensor T_Ω, suppose that if there are two or more of i, j, or k less than or equal to r, then $(i, j, k) \in \Omega$. Let A denote the $r \times r \times r$ known subtensor of T_Ω consisting of entries in positions (i, j, k) where $i \leq r$, $j \leq r$, and $k \leq r$. Suppose A has multilinear rank equal to (r, r, r). Also, suppose $\mathcal{A}_\Omega \cap \hat{S}ub_r$ is non-empty. Then T_Ω has a unique multilinear rank (r, r, r) completion $T \in \mathcal{A}_\Omega \cap \hat{S}ub_r$. Moreover, if $\mathcal{A}_\Omega \cap \sigma_r$ or $\mathcal{A}_\Omega \cap \hat{\sigma}_r$ are non-empty, then T is also the unique rank r or border rank r completion of T_Ω respectively.*

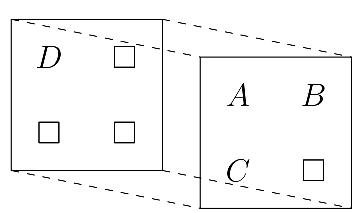


Figure 4.3: Incomplete tensor T_Ω, with known subtensors A, B, C, D

Proof. Because $\hat{S}ub_r = \{T \in U \otimes V \otimes W \mid \text{rank}(T^{(i)}) \leq r, i = 1, 2, 3\}$, the equations that generate $\hat{S}ub_r$ are all $(r+1) \times (r+1)$ minors of each mode-i unfolding. In other words, a tensor T is in $\hat{S}ub_r$ if and only if all $(r+1) \times (r+1)$ minors vanish in each of the unfoldings

$$T^{(1)} : (V \otimes W)^* \to U$$
$$T^{(2)} : (U \otimes W)^* \to V$$
$$T^{(3)} : (U \otimes V)^* \to W.$$

Let T_Ω be given as defined in Theorem 117. Let B denote the known subtensor of T_Ω consisting of entries with positions (i, j, k) such that $i \leq r$, $j > r$, and $k \leq r$. Let C denote the known subtensor of T_Ω consisting of entries with positions (i, j, k) such that $i > r$, $j \leq r$, and $k \leq r$. Let D denote the known subtensor of T_Ω consisting of entries with positions (i, j, k) such that $i \leq r$, $j \leq r$, and $k > r$.

First, we show that we may complete entries T_{ijk} where exactly one of j or k is less than or equal to r by using the mode-1 unfolding. Then, we complete the rest of the entries by using the mode-2 unfolding. Because the known subtensor A has multilinear rank (r, r, r), the mode-1 unfolding $A^{(1)}$ has at least one rank r submatrix. Choose a rank r submatrix in $A^{(1)}$, and denote it A_J. Define $J = \{(j_\alpha, k_\alpha)\}_{1 \leq \alpha \leq r}$ as the set of r pairs of indices such that the entry in position (i, α) of A_J is equal to $T_{ij_\alpha k_\alpha}$.

Let G denote the subtensor of T with entries T_{ijk} such that $k \leq r$, $i > r$, and $j > r$. Then each entry T_{ijk} of G is unknown, and there is a $(r+1) \times (r+1)$ submatrix of the mode-1 flattening of T_Ω of the form

$$\begin{bmatrix} A_J & b_{jk} \\ c_{iJ} & T_{ijk} \end{bmatrix}$$

where $b_{jk} = [T_{ljk}]_{1 \leq l \leq r}$ is the $r \times 1$ submatrix of $B^{(1)}$ consisting of entries of the form T_{ljk}, with j and k fixed, and l ranging from 1 to r. Also $c_{iJ} = [T_{ij_\alpha k_\alpha}]_{1 \leq \alpha \leq r}$ is the $1 \times r$ submatrix of $C^{(1)}$ where i is fixed, and $(j_\alpha, k_\alpha) \in J$ with α ranging from 1 to r. Because every $(r+1) \times (r+1)$ minor of $T^{(1)}$ must vanish, and because A_J is invertible, we may set the determinant of this submatrix equal to zero and solve for T_{ijk}, getting $T_{ijk} = c_{iJ} A_J^{-1} b_{jk}$, which completes G.

Let E denote the subtensor of T with entries T_{ijk} such that $j \leq r$, $i > r$, and $k > r$. Then T_{ijk} is unknown, and there is a $(r+1) \times (r+1)$ submatrix of the mode-1 flattening of T_Ω of the form

$$\begin{bmatrix} A_J & d_{jk} \\ c_{iJ} & T_{ijk} \end{bmatrix}$$

where $d_{jk} = [T_{ljk}]_{1 \leq l \leq r}$ is the $r \times 1$ submatrix of $D^{(1)}$ consisting of entries of the form T_{ljk}, with j and k fixed, and l ranging from 1 to r. Again, setting the determinant of this submatrix equal to zero we may solve for T_{ijk}, getting $T_{ijk} = c_{iJ} A_J^{-1} d_{jk}$, which completes E.

Now we consider the mode-2 unfolding of T_Ω. Again, because A has multilinear rank (r, r, r), the mode-2 unfolding $A^{(2)}$ has at least one rank r submatrix. Choose a rank r submatrix of $A^{(2)}$, and denote it A_I. Define $I = \{(i_\beta, k_\beta)\}_{1 \leq \beta \leq r}$ as the set of r pairs of indices such that the entry in position (j, β) of A_I is equal to $T_{i_\beta j k_\beta}$.

Let F denote the subtensor of T with entries T_{ijk} such that $i \leq r$, $j > r$, and $k > r$. Then each entry T_{ijk} of F is unknown, and there is a $(r+1) \times (r+1)$ submatrix of the mode-2 flattening of T_Ω of the form

$$\begin{bmatrix} A_I & d_{ik} \\ b_{jI} & T_{ijk} \end{bmatrix}$$

where $d_{ik} = [T_{ilk}]_{1 \leq l \leq r}$ is the $r \times 1$ submatrix of $D^{(2)}$ consisting of entries of the form T_{ilk}, with i and k fixed, and l ranging from 1 to r. Also $b_{jI} = \left[T_{i_\beta j k_\beta}\right]_{1 \leq \beta \leq r}$ is the $1 \times r$ submatrix of $B^{(2)}$ where j is fixed, and $(i_\beta, k_\beta) \in I$ with β ranging from 1 to r. Because every $(r+1) \times (r+1)$ minor of $T^{(2)}$ must vanish, and because A_I is invertible, we may set the determinant of this submatrix equal to zero and solve for T_{ijk}, getting $T_{ijk} = b_{jI} A_I^{-1} d_{ik}$ which completes F.

Finally, let H denote the subtensor of T with entries T_{ijk} such that $i > r$, $j > r$, and $k > r$. Then each entry T_{ijk} of H is unknown, and there is a $(r+1) \times (r+1)$ submatrix of the mode-2 flattening of

T_Ω of the form

$$\begin{bmatrix} A_I & e_{ik} \\ b_{jI} & T_{ijk} \end{bmatrix}$$

where $e_{ik} = [T_{ilk}]_{1 \leq l \leq r}$ is the $r \times 1$ submatrix of $E^{(1)}$ consisting of entries of the form T_{ilk}, with i and k fixed, and l ranging from 1 to r. Again setting the determinant of this submatrix equal to zero we may solve for T_{ijk}, getting $T_{ijk} = b_{jI} A_I^{-1} e_{ik}$, which completes H, and finishes completing T_Ω.

Note that every tensor in $\mathcal{A}_\Omega \cap \hat{S}ub_r$ must be in the zero set of the system of equations used to construct the completion T. Because each equation had a unique solution, each unknown entry is uniquely determined, and because $\mathcal{A}_\Omega \cap \hat{S}ub_r$ is non-empty, the constructed completion T is unique.

Moreover, because $T \in \hat{S}ub_r$, each component of the multilinear rank is at most r. Also, because A is a subtensor of T, and $\mu\text{rank}(A) = (r, r, r)$, then each component of the multilinear rank of T is at least r, so $\mu\text{rank}(T) = (r, r, r)$

Also note that by Theorem 114, if $\mathcal{A}_\Omega \cap \sigma_r$ is non-empty, then T is a rank at most r completion of T_Ω, and because $\max(\mu\text{rank}(T)) = r \leq \text{rank}(T) \leq r$, the rank of T is equal to r. Similarly, if $\mathcal{A}_\Omega \cap \hat{\sigma}_r$ is non-empty, then T is a border rank r completion of T_Ω. □

In total $r^2(m+n+p) - r^3$ of mnp entries are known. This is an improvement from (Cai et al., 2020) in which $O(nr^2 + r^4)$ known entries are required when $m = n = p$. We have the following corollary.

Corollary 118. *Given an $m \times n \times p$ partially known tensor T_Ω, if the positions of the known entries in T_Ω are distributed correctly, and $\mathcal{A}_\Omega \cap \hat{S}ub_r$ is non-empty, then $r^2(n+m+p) - r^3$ observed entries is a sufficient condition for T_Ω to have a unique completion in $\hat{S}ub_r$.*

Note there were entries in A that we did not use to compute unknown entries, we only used entries in A_J and A_I. Moreover, we did not use the fact that the mode-3 unfolding of A is rank r, so it is likely possible to express a similar unique tensor completion theorem under weaker assumptions.

Recall that Example 32 is the matrix analog of Theorem 117, in which case the existence of a rank r completion always exits under the assumption that A is full rank. In fact, a rank r completion of M_Ω

will exist with probability one if the entries are sampled from a continuous density. It may be possible that observing a fewer number of entries would guarantee the existence of a multilinear rank (r, r, r) completion of $T_\Omega = P_\Omega(T)$ with probability one.

We now construct an explicit multilinear rank $(2, 2, 2)$ completion for an example T_Ω under the assumptions of Theorem 117.

Example 119. *Consider the partially known $3 \times 3 \times 3$ tensor T_Ω from Figure 4.4. Note that any way to unfold T_Ω results in an incomplete row or column. Therefore unfolding T_Ω in one way is not sufficient to complete T_Ω, we must use multiple different unfoldings of T_Ω to find a rank $(2, 2, 2)$ completion.*

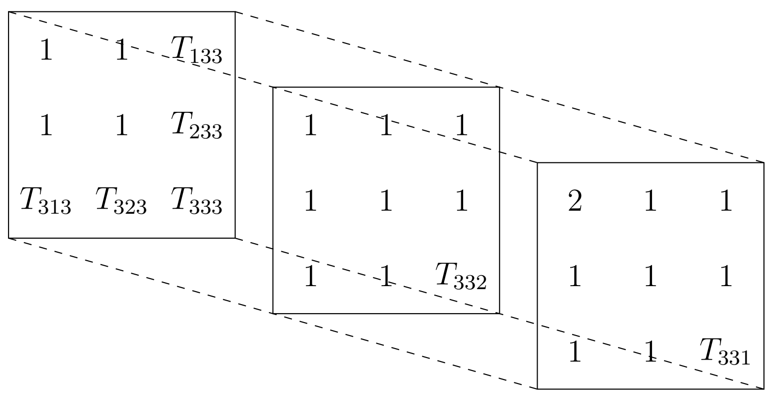


Figure 4.4: $3 \times 3 \times 3$ incomplete tensor T_Ω from Example 119 with unknown entries T_{ijk}

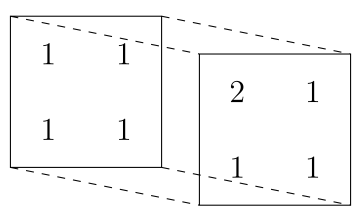


Figure 4.5: $2 \times 2 \times 2$ known subtensor A of T_Ω with multilinear rank $(2, 2, 2)$

We construct a multilinear rank $(2, 2, 2)$ completion T of T_Ω. Note that the subtensor, $A = [T_{ijk}]$ with i, j, and k at most two, is fully known and has multilinear rank $(2, 2, 2)$. Moreover, T_{ijk} is known if two or more of i, j, or k are at most two. Therefore, we may use Theorem 117 to attempt to construct a unique multilinear rank $(2, 2, 2)$ completion of T_Ω.

Consider the mode-1 *unfoldings of* T_Ω *and* A

$$T_\Omega^{(1)} = \left[\begin{array}{ccc|ccc|ccc} 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & T_{133} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & T_{233} \\ 1 & 1 & T_{331} & 1 & 1 & T_{332} & T_{313} & T_{323} & T_{333} \end{array}\right]$$

$$A^{(1)} = \left[\begin{array}{cc|cc} 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}\right].$$

First, we recover $T_{331}, T_{332}, T_{313}$*, and* T_{323} *by solving the equations of the form*

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & T_{ijk} \end{vmatrix} = 0 \tag{4.2}$$

getting $T_{ijk} = 1$*. Next, we consider the mode-*2 *unfolding of* T_Ω*,*

$$T_\Omega^{(2)} = \left[\begin{array}{ccc|ccc|ccc} 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & T_{313} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & T_{323} \\ 1 & 1 & T_{133} & 1 & 1 & T_{233} & T_{331} & T_{332} & T_{333} \end{array}\right].$$

Filling in the completed entries from the mode-1 *unfolding, we get*

$$T^{(2)} = \left[\begin{array}{ccc|ccc|ccc} 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & T_{133} & 1 & 1 & T_{233} & 1 & 1 & T_{333} \end{array}\right].$$

We may now recover entries T_{133}, T_{233}, *and* T_{333} *like Equation* (4.2). *There also exists a rank two completion of* T_Ω, *and so the completed tensor* T *is also rank two, and border rank two. That is, we have*

$$T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

If the slices of T_Ω may be permuted to the appropriate form, Theorem 117 gives a constructive way to complete T_Ω to a multilinear rank (r, r, r) completion, assuming one exists. We now explicitly express the tensor completion algorithm. Recall that in the proof of Theorem 117, T is partitioned into the subtensors A, B, C, D, E, F, G, and H.

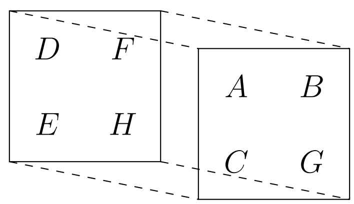


Figure 4.6: T partitioned into eight subtensors where A is $r \times r \times r$

Subtensor of T	Entries T_{ijk} with
A	$i \leq r, j \leq r, k \leq r$
B	$i \leq r, j > r, k \leq r$
C	$i > r, j \leq r, k \leq r$
D	$i \leq r, j \leq r, k > r$
E	$i > r, j \leq r, k > r$
F	$i \leq r, j > r, k > r$
G	$i > r, j > r, k \leq r$
H	$i > r, j > r, k > r$

Then if T_Ω satisfies the assumptions given in Theorem 117, we may use Algorithm 15 to complete T_Ω into a multilinear rank (r, r, r) tensor.

Algorithm 15: Exact Tensor Completion

Input: partially known tensor T_Ω satisfying assumptions from Theorem 117
Result: completed tensor T such that $\mu\text{rank}(T) = (r, r, r)$ and $P_\Omega(T) = T_\Omega$

find a full rank $r \times r$ submatrix A_J of $A^{(1)}$;
$J = \{(j_\alpha, k_\alpha)\}_{1 \leq \alpha \leq r}$ is the set of r pairs of indices such that the entry in position (i, α) of A_J is equal to $T_{ij_\alpha k_\alpha}$;
C_J equals the $(m - r) \times r$ submatrix of $C^{(1)}$ such that the entry in position (i, α) of C_J equals the entry of T in position $(i + r, j_\alpha, k_\alpha)$;
$G^{(1)} = C_J A_J^{-1} B^{(1)}$;
$E^{(1)} = C_J A_J^{-1} D^{(1)}$;
reshape $E^{(1)}$ and $G^{(1)}$ obtaining E and G;
find a full rank $r \times r$ submatrix A_I of $A^{(2)}$;
$I = \{(i_\beta, k_\beta)\}_{1 \leq \beta \leq r}$ is the set of r pairs of indices such that the entry in position (j, β) of A_I is equal to $T_{i_\beta j k_\beta}$;
B_I equals the $(n - r) \times r$ submatrix of $B^{(2)}$ such that the entry in position (j, β) of B_I equals the entry of T in position $(i_\beta, j + r, k_\beta)$;
$F^{(2)} = B_I A_I^{-1} D^{(2)}$;
$H^{(2)} = B_I A_I^{-1} E^{(2)}$;
reshape $F^{(2)}$ and $H^{(2)}$ obtaining F and H;
assemble subtensors, output completed T;

In practice T_Ω may not have an exact multilinear rank r completion, so one may choose A_J and A_I from $A^{(1)}$ and $A^{(2)}$ respectively with a maximum volume algorithm such as Algorithm 8 or Algorithm 12 to optimize the quality of the completion.

We conjecture that analogous algorithms work to complete higher order tensors with similar structures.

Conjecture 120. *Given an order $d > 3$ partially known $n_1 \times \cdots \times n_d$ tensor T_Ω, suppose that given a position $(i_1, \ldots, i_d)$, if there are $d - 1$ or more entries less than or equal to r, then $(i_1, \ldots, i_d)$ is in Ω. Suppose the known subtensor $A = [T_{i_1 \cdots i_d}], i_j \leq r$ for all j has multilinear rank equal to $(r, \ldots, r)$, and suppose also that there exists at least one multilinear rank $(r, \ldots, r)$ completion of T_Ω. Then T_Ω has a unique multilinear rank $(r, \ldots, r)$ completion.*

This means that in total $(\sum_{i=1}^{d} n_i) r^{d-1} - (d-1) r^d$ entries are known of $\prod_{i=1}^{d} n_i$ entries total. In this case, the expected way to recover T from T_Ω is to first complete entries in positions of the form $(i_1, \ldots, i_d)$

where at least $d-2$ entries are at most r, then recover entries in positions of the form $(i_1,\ldots,i_d)$ where at least $d-3$ entries are at most r, and so on.

4.3 Algebraic Combinatorics

In this section we will generalize a few concepts from Section 1.10 for tensors.

Definition 121. *A 3-partite hypergraph is a tuple (V_1,V_2,V_3,E) where V_1, V_2, and V_3 are the sets of vertices, and $E\subset V_1\times V_2\times V_3$ are the hyperedges which consist of one vertex from each of V_1, V_2, and V_3.*

Given a mask $\Omega\subset[m]\times[n]\times[p]$, define $G_\Omega=(V_1,V_2,V_3,\Omega)$ as the 3-partite hypergraph with vertices $V_1=[m]$, $V_2=[n]$, $V_3=[p]$, and hyperedges Ω. There is a one to one correspondence between masks of third order tensors and 3-partite hypergraphs.

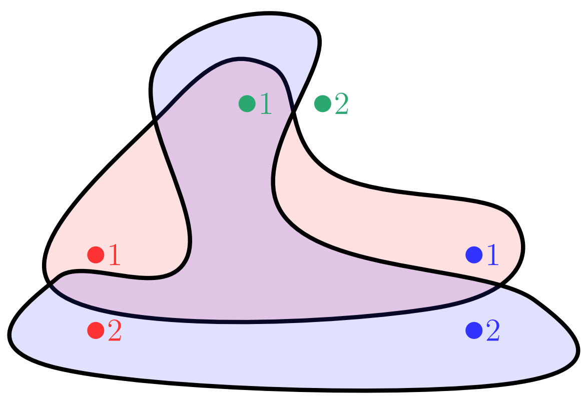


Figure 4.7: The 3-partite hypergraph of a $2\times2\times2$ mask Ω. Two hyperedges correspond to positions $(1,1,1)$ and $(2,1,2)$.

Definition 122. *Given a mask Ω, a position $(i,j,k)\in[m]\times[n]\times[p]$ is called* finitely completable in rank r *if, for a generic $T\in\hat{S}ub_r$, the entry in position (i,j,k) of the partially known tensor $P_\Omega(T)$ has finitely many maximum multilinear rank at most r completions. Define the* rank r finitely completable closure $\mathrm{cl}_r(\Omega)$ *as the set of positions which are finitely completable in Ω.*

In terms of the rank r finitely completable closure $\mathrm{cl}_r(\Omega)$, we may reformulate Theorem 117 as follows.

Theorem 123. *Given Ω, suppose if there are two or more entries of (i, j, k) less than or equal to r, then $(i, j, k) \in \Omega$. Then $\mathrm{cl}(\Omega)$ is equal to $[m] \times [n] \times [p]$.*

Proof. Suppose (i, j, k) having two or more indices less than or equal to r implies $(i, j, k) \in \Omega$. We have shown in Theorem 117 that for $T \in \hat{S}ub_r$, if the subtensor $A = [T_{ijk}]\,, 1 \leq i, j, k \leq r$ has multilinear rank equal to (r, r, r), then $P_\Omega(T)$ can be uniquely completed to T. Note that for a generic T, A is also generic, which implies any mode-i unfolding $A^{(i)}$ of A is full rank r. So $A^{(i)}$ will contain a rank r submatrix almost surely for all i. Therefore, the subtensor A will have multilinear rank (r, r, r) almost surely. So a generic tensor T will have subtensor A with multilinear rank equal to (r, r, r), which means $P_\Omega(T)$ can be uniquely completed to T, and so every entry is finitely completable. □

We may also generalize the unique completability closure to the case of tensors.

Definition 124. *Given a mask Ω, a position $(i, j, k) \in [m] \times [n] \times [p]$ is called* uniquely completable in rank r *if, for a generic $T \in \hat{S}ub_r$, the entry in position (i, j, k) of the partially known tensor $P_\Omega(T)$ has a unique maximum multilinear rank at most r completions. Define the* rank r uniquely completable closure $\mathrm{ucl}_r(\Omega)$ *as the set of positions which are finitely completable in Ω.*

By Theorem 117, the finitely completable closure equals the uniquely completable closure which is equal to all of $[m] \times [n] \times [p]$

Theorem 125. *Given Ω as defined in Theorem 117, then we have* $\mathrm{ucl}_r(\Omega) = \mathrm{cl}_r(\Omega) = [m] \times [n] \times [p]$.

Proof. The result follows from Theorem 123 and the fact that Theorem 117 gives a unique completion. □

4.4 Algebraic Topology

We may generalize many of the algebraic topology results for low-rank matrix completion from Section 1.11 to low-rank tensor completion. The spaces $\hat{\sigma}_r$, $\hat{S}ub_r$, and $\overline{\mathcal{M}_r}$ are alike in the sense that they are affine algebraic varieties which deformation retract to a point. In particular, the border rank and multilinear rank are invariant under scaling (Landsberg, 2011). That is, given a tensor T, for any non-zero constant α,

$$\overline{\text{rank}}(\alpha T) = \overline{\text{rank}}(T)$$
$$\mu\text{rank}(\alpha T) = \mu\text{rank}(T).$$

This implies that $\hat{\sigma}_r$ and $\hat{S}ub_r$ deformation retract to $\{0\}$ through the map

$$F : V \times [0,1] \to V$$
$$F(T,\alpha) = (1-\alpha)T$$

where V is $\hat{\sigma}_r$ or $\hat{S}ub_r$. Similarly, $\mathcal{A}_\Omega$ is a linear variety so it deformation retracts to the singleton $\{T_\Omega\}$ via the map

$$G : \mathcal{A}_\Omega \times [0,1] \to \mathcal{A}_\Omega$$
$$G(T,\alpha) = (1-\alpha)T + \alpha T_\Omega.$$

Both $\hat{\sigma}_r$ and $\hat{S}ub_r$ are affine algebraic varieties, and so may be simultaneously triangulated with $\mathcal{A}_\Omega$ by Theorem 55. So we have the following theorem.

Theorem 126. *Given a partially known tensor T_Ω, if T_Ω has at least one completion in $\hat{\sigma}_r$, then*

$$h_0(\mathcal{A}_\Omega \cap \hat{\sigma}_r) = h_1(\mathcal{A}_\Omega \cup \hat{\sigma}_r) + 1$$
$$\chi(\mathcal{A}_\Omega \cap \hat{\sigma}_r) = 2 - \chi(\mathcal{A}_\Omega \cup \hat{\sigma}_r).$$

If T_Ω has at least one completion in $\hat{S}ub_r$, then

$$h_0(\mathcal{A}_\Omega \cap \hat{S}ub_r) = h_1(\mathcal{A}_\Omega \cup \hat{S}ub_r) + 1$$
$$\chi(\mathcal{A}_\Omega \cap \hat{S}ub_r) = 2 - \chi(\mathcal{A}_\Omega \cup \hat{S}ub_r).$$

Proof. The proof is identical to the proofs of Theorem 56 and Theorem 61. Let V be $\hat{\sigma}_r$ or $\hat{S}ub_r$. Suppose T_Ω has at least one completion in V, then $\mathcal{A}_\Omega \cap V$ is nonempty. By Theorem 54 and Theorem 55, we have an exact sequence of the form

$$\cdots \longrightarrow H_1(\mathcal{A}_\Omega) \oplus H_1(V) \longrightarrow H_1(\mathcal{A}_\Omega \cup V) \longrightarrow H_0(\mathcal{A}_\Omega \cap V) \longrightarrow H_0(\mathcal{A}_\Omega) \oplus H_0(V) \longrightarrow H_0(\mathcal{A}_\Omega \cup V) \longrightarrow 0$$

Because homology groups are invariant under homotopy equivalence, V and $\overline{\mathcal{M}}_r$ have the homology of a point. So

$$\begin{aligned} H_1(V) &\cong 0 & H_1(\mathcal{A}_\Omega) &\cong 0 \\ H_0(V) &\cong \mathbb{Z} & H_0(\mathcal{A}_\Omega) &\cong \mathbb{Z} \\ \chi(V) &= 1 & \chi(\mathcal{A}_\Omega) &= 1. \end{aligned}$$

Because $\mathcal{A}_\Omega$ and V are both connected and have non-empty intersection, their union is also connected, so

$$H_0(\mathcal{A}_\Omega \cup V) \cong \mathbb{Z}.$$

So we have

$$0 \longrightarrow H_1(\mathcal{A}_\Omega \cup V) \longrightarrow H_0(\mathcal{A}_\Omega \cap V) \longrightarrow \mathbb{Z}^2 \longrightarrow \mathbb{Z} \longrightarrow 0. \tag{4.3}$$

After tensoring with $\mathbb{R}$, the alternating sum of the vector spaces vanish, so we have

$$h_0(\mathcal{A}_\Omega \cap V) = h_1(\mathcal{A}_\Omega \cup V) + 1.$$

Also applying Theorem 60, we have

$$\chi(\mathcal{A}_\Omega \cap V) = 2 - \chi(\mathcal{A}_\Omega \cup V).$$

$\square$

If T_Ω has $N < \infty$ completions in V where V is $\hat{\sigma}_r$ or $\hat{S}ub_r$, then

$$h_0(\mathcal{A}_\Omega \cap V) = \chi(\mathcal{A}_\Omega \cap V) = N.$$

So

$$N = h_1(\mathcal{A}_\Omega \cup V) + 1 = 2 - \chi(\mathcal{A}_\Omega \cup V).$$

4.5 Maximum Volume Gradient Descent for Tensor Completion

We now generalize the Schur gradient descent method from Section 2.7 to the case of order three tensors. Let T_Ω be a partially known $m \times n \times p$ tensor. The objective is to find a tensor T such that $P_\Omega(T) = T_\Omega$ and $\max(\mu\text{rank}(T)) \leq r$, which we cast as a minimization problem.

Recall that $T \in U \otimes V \otimes W$ has the mode-1, mode-2, and mode-3 unfoldings

$$T^{(1)} : (V \otimes W)^* \to U$$
$$T^{(2)} : (U \otimes W)^* \to V$$
$$T^{(3)} : (U \otimes V)^* \to W.$$

Let A_1, A_2, and A_3 denote the top-left $r \times r$ submatrices of $T^{(1)}$, $T^{(2)}$, and $T^{(3)}$ respectively. To recover missing elements from the partially known tensor T_Ω, we may attempt to solve the minimization problem

$$\min_T \left(\frac{1}{2} \|S_{A_1}\|^2 + \frac{1}{2} \|S_{A_2}\|^2 + \frac{1}{2} \|S_{A_3}\|^2 \right) \tag{4.4}$$
$$\textit{s.t. } P_\Omega(T) = T_\Omega$$

where S_{A_i} denotes the Schur complement of $T^{(i)}$ with respect to A_i. Let

$$f(T) = \frac{1}{2} \left(\|S_{A_1}\|^2 + \|S_{A_2}\|^2 + \|S_{A_3}\|^2 \right) \tag{4.5}$$

Then $f(T)$ is non-negative because it is the sum of squares of norms. So if $f(T) = 0$ and $P_\Omega(T) = T_\Omega$, then T is a solution to Minimization (4.4). In particular, we have the following theorem.

Theorem 127. *$f(T) = 0$ if and only if $\mu\text{rank}(T) = (r, r, r)$ and A_1, A_2, and A_3 are full rank r.*

Proof. Suppose $f(T) = 0$. Then $\|S_{A_i}\| = 0$ for all i, so $S_{A_i} = 0$ for all i. So $\text{rank}(T^{(i)}) = r$ for all i and $\mu\text{rank}(T) = (r, r, r)$ by lemma 66.

On the other hand, if $\mu\text{rank}(T) = (r, r, r)$ and A_1, A_2, and A_3 are full rank r, then $S_{A_i} = 0$ for all i by lemma 66, so $f(T) = 0$. □

To find a minimizer T, we may employ a gradient descent method like in Section 2.7. In particular, we have the following derivative information.

Theorem 128. *Given an $m \times n \times p$ tensor T, let*

$$T^{(1)} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}$$
$$T^{(2)} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}$$
$$T^{(3)} = \begin{bmatrix} A_3 & B_3 \\ C_3 & D_3 \end{bmatrix}$$

be block matrices where A_i is the top-left $r \times r$ submatrix of $T^{(i)}$, and B_i, C_i, and D_i are corresponding block submatrices. Then we have

$$\partial f = \sum_{i=1}^{3} \langle S_{A_i}, \partial D_i \rangle + \langle A_i^{-\top} C_i^\top S_{A_i} B_i^\top A_i^{-\top}, \partial A_i \rangle - \langle S_{A_i} B_i^\top A_i^{-\top}, \partial C_i \rangle - \langle A_i^{-\top} C_i^\top S_{A_i}, \partial B_i \rangle$$

Proof. The result follows from the identities given in Proposition 95 as in the proof of Theorem 96. □

Note that it may be tricky to calculate derivatives such as $\frac{\partial D_1}{\partial D_2}$ because A_i, B_i, C_i, and D_i are functions of A_j, B_j, C_j, and D_j when $i \neq j$.

Like Algorithm 14, we alternate between taking gradient descent steps and applying a maximum volume algorithms such as Algorithm 10 or an alternating version of Algorithm 12 to each unfolding. Given a function $f : U \otimes V \otimes W \to \mathbb{R}$ such as Equation (4.5), let $\frac{\partial f}{\partial T}$ denote the order three tensor such that

$$\left(\frac{\partial f}{\partial T}\right)_{ijk} = \frac{\partial f}{\partial T_{ijk}}.$$

Using the derivative information, Algorithm 16 as a generalization of Algorithm 14 for tensor completion.

We apply P_{Ω^c} to the gradient so the known entries of X_k remain unchanged at each iteration.

Algorithm 16: Maximum Volume Gradient Descent for Tensor Completion

Input: partially known tensor T_Ω, initial guess X_0 such that $P_\Omega(X_0) = T_\Omega$, $r \times r$ nonsingular submatrices $A_{1,0}$ of $X_0^{(1)}$, $A_{2,0}$ of $X_0^{(2)}$, and $A_{3,0}$ of $X_0^{(3)}$, sequence of step sizes $\{h_k\}$, tolerance $\epsilon > 0$, stopping criterion

Result: X_k an approximation of a multilinear rank (r, r, r) completion of T_Ω

for $k = 1, \ldots$ **do**

$A_{i,k}$ are submatrices of $X_{k-1}^{(i)}$, $i = 1, 2, 3$, which are the outputs of a two directional maximum volume algorithm, such as Algorithm 10 or an alternating version of Algorithm 12, with initialization $X_{k-1}^{(i)}$, initial nonsingular submatrix $A_{i,k-1}$, tolerance ϵ;

$f_k(X_{k-1}) = \frac{1}{2}\left(\left\|S_{A_{1,k}}\right\|^2 + \left\|S_{A_{2,k}}\right\|^2 + \left\|S_{A_{3,k}}\right\|^2\right)$;

$X_k = X_{k-1} - h_k P_{\Omega^c}\left(\frac{\partial f_k(X_{k-1})}{\partial T}\right)$;

Chapter 5

Conclusions

We have introduced novel theoretical results on low-rank matrix and tensor completion. We have shown in Section 1.9 that we may reduce the number of polynomials needed to generate the space of low-rank completions of a partially known matrix by finding a known full rank submatrix. A remaining question is to determine if the generating set of polynomials can be reduced even further.

In Section 1.11 and Section 4.4 we discuss some topological properties of the space of low-rank completions of a partially known matrix or tensor. We relate the zeroth Betti number and Euler characteristic of the space of low-rank completions to the first Betti number and Euler characteristic respectively of the union of the space of low-rank matrices or tensors and the space of all completions.

We propose a method to reduce the processing time of the maximum volume algorithm, Algorithm 8, in Section 2.4 by potentially permuting more than one row at each step. We call such an algorithm a "greedy" maximum volume algorithm which is presented in Algorithm 12. We give numerical evidence to show that the greedy maximum volume algorithm may reduce the processing time needed for the maximum volume algorithm.

A novel connection is made between dominant submatrices and Johnson graphs in Section 2.6. In particular, we show that the independence number of the graph Cartesian product of Johnson graphs given an upper bound on the number of $r \times r$ dominant submatrices in almost every matrix. An open question is to determine if this independence number is the essential supremum of the function which counts the number of dominant submatrices, which means the inequality is sharp on a set of positive measure.

We introduce a novel approach to low-rank matrix completion in Section 2.7 which utilizes the maximum volume skeleton decomposition. One of the benefits of this method is that it does not utilize the singular value decomposition (SVD) which scales poorly as the matrix gets large. It may be possible to improve this method with more sophisticated gradient descent techniques. In Section 4.5 we propose a way to generalize this method to low-rank tensor completion.

One of our primary goals in Section 3.2 was to test if the maximum volume skeleton decomposition could be a suitable substitute for the singular value decomposition as a low-rank approximation step in the dynamic mode decomposition. We have shown that the maximum volume skeleton decomposition (MVSD) can be a reasonable direct substitute for the SVD as a low-rank approximation when applied low-noise temperature data. For theoretical results, we have reproduced several theorems by (Tu, 2013) using the MVSD as a substitute for the SVD. We have shown that certain desirable properties hold when using the MVSD as a low-rank approximation. However, the full-matrix is only a rank r approximation of the original matrix. The conditions which make that approximation a suitable substitute for the full matrix are not entirely known, but are important when generating the dynamics.

We have also studied applications of the MVSD for compressing simulation data in Section 3.3. We have shown that if a certain threshold of error ϵ is acceptable, the MVSD may be multiple orders of magnitude faster than the SVD for finding a low-rank approximation with error ϵ. We compare low-rank approximations found by the MVSD and the SVD when applied to two dimensional in time temperature and density data unfolded into a 1633×3724 matrix. Moreover, we have shown that an alternating version of the greedy maximum volume algorithm, Algorithm 12, may further reduce the processing time while fixing the error.

Low-rank approximations used for compression or in the dynamic mode decomposition are often generated using the SVD. When the size of the matrix is large, the SVD can be computationally prohibitive. We have shown that a maximum volume skeleton decomposition such as Algorithm 10 can significantly reduce the processing time needed to find a low-rank approximation of numerical simulation data at the cost of moderately increased error. If the acceptable error threshold is large enough, the MVSD may be a desirable substitute for the SVD.

Finally, we introduce a class of partially known tensors which have unique low-rank, low-border-rank, or low-multilinear-rank completions under the assumption that one exists in Section 4.2. We also give an explicit algorithm to construct such a completion. It may be possible to find a more general class of partially known tensors which have unique low-rank completions by weakening the assumptions on the structure partially known tensor.

Bibliography

Aliprantis, C. D., & Burkinshaw, O. (1998). *Principles of real analysis*. Gulf Professional Publishing.

Bonnin, X., Dekeyser, W., Pitts, R., Coster, D., Voskoboynikov, S., & Wiesen, S. (2016). Presentation of the new solps-iter code package for tokamak plasma edge modelling. *Plasma and Fusion Research*, *11*, 1403102–1403102.

Brouwer, A. E., & Etzion, T. (2011). Some new distance-4 constant weight codes. *Advances in Mathematics of Communications*, *5*(3), 417.

Brouwer, A. E., Shearer, J. B., Sloane, N. J., & Smith, W. D. (2006). A new table of constant weight codes. *IEEE Transactions on Information Theory*, *36*(6), 1334–1380.

Burnside, W. (1911). *Theory of groups of finite order*. University Press.

Cai, J.-F., Candès, E. J., & Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. *SIAM Journal on optimization*, *20*(4), 1956–1982.

Cai, J.-F., Miao, L., Wang, Y., & Xian, Y. (2020). Provable near-optimal low-multilinear-rank tensor recovery. *arXiv preprint arXiv:2007.08904*.

Candès, E. J., & Tao, T. (2010). The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, *56*(5), 2053–2080.

Civril, A., & Magdon-Ismail, M. (2009). On selecting a maximum volume sub-matrix of a matrix and related problems. *Theoretical Computer Science*, *410*(47-49), 4801–4811.

De Silva, V., & Lim, L.-H. (2008). Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, *30*(3), 1084–1127.

del-Castillo-Negrete, D., Hirshman, S. P., Spong, D. A., & D'Azevedo, E. F. (2007). Compression of magnetohydrodynamic simulation data using singular value decomposition. *Journal of Computational Physics*, *222*(1), 265–286.

Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*(3), 211–218.

Edelman, A. (1988). Eigenvalues and condition numbers of random matrices. *SIAM journal on matrix analysis and applications*, *9*(4), 543–560.

Fazel, M. (2002). *Matrix rank minimization with applications* (Doctoral dissertation). PhD thesis, Stanford University.

Federici, G., Skinner, C. H., Brooks, J. N., Coad, J. P., Grisolia, C., Haasz, A. A., Hassanein, A., Philipps, V., Pitcher, C. S., Roth, J., Et al. (2001). Plasma-material interactions in current tokamaks and their implications for next step fusion reactors. *Nuclear Fusion*, *41*(12), 1967.

Friedland, S. (2012). On the generic and typical ranks of 3-tensors. *Linear algebra and its applications*, *436*(3), 478–497.

Fulton, W. (2016). *Intersection theory*. Princeton University Press.

Goreinov, S. A., Oseledets, I. V., Savostyanov, D. V., Tyrtyshnikov, E. E., & Zamarashkin, N. L. (2010). How to find a good submatrix, In *Matrix methods: Theory, algorithms and applications: Dedicated to the memory of gene golub*. World Scientific.

Guillemin, V., & Pollack, A. (2010). *Differential topology* (Vol. 370). American Mathematical Soc.

Hadamard, J. (1893). Resolution d'une question relative aux determinants. *Bull. des sciences math.*, *2*, 240–246.

Harris, J., & Tu, L. W. (1984). On symmetric and skew-symmetric determinantal varieties.

Hartshorne, R. (2013). *Algebraic geometry* (Vol. 52). Springer Science & Business Media.

Hatch, D. R., del-Castillo-Negrete, D., & Terry, P. W. (2012). Analysis and compression of six-dimensional gyrokinetic datasets using higher order singular value decomposition. *Journal of Computational Physics*, *231*(11), 4234–4256.

Hatcher, A. (2002). *Algebraic topology*. Cambridge University Press.

Hironaka, H. (1975). Triangulations of algebraic sets, In *Algebraic geometry (proc. sympos. pure math., vol. 29, humboldt state univ., arcata, calif., 1974)*.

Jain, P., Netrapalli, P., & Sanghavi, S. (2013). Low-rank matrix completion using alternating minimization, In *Proceedings of the forty-fifth annual acm symposium on theory of computing*.

Johnson, S. (1962). A new upper bound for error-correcting codes. *IRE Transactions on Information Theory*, *8*(3), 203–207.

Kaveeva, E., Rozhansky, V., Senichenkov, I., Veselova, I., Voskoboynikov, S., Sytova, E., Bonnin, X., & Coster, D. (2018). Speed-up of solps-iter code for tokamak edge modeling. *Nuclear Fusion*, *58*(12), 126018.

Király, F. J., Theran, L., & Tomioka, R. (2015). The algebraic combinatorial approach for low-rank matrix completion. *J. Mach. Learn. Res.*, *16*(1), 1391–1436.

Lai, M. J., & Varghese, A. (2017). On convergence of the alternating projection method for matrix completion and sparse recovery problems. *arXiv preprint arXiv:1711.02151*.

Landsberg, J. (2011). *Tensors: Geometry and applications: Geometry and applications* (Vol. 128). American Mathematical Soc.

Landsberg, J., & Michałek, M. (2017). Abelian tensors. *Journal de Mathématiques Pures et Appliquées*, *108*(3), 333–371.

Lewis, A. S., & Malick, J. (2008). Alternating projections on manifolds. *Mathematics of Operations Research*, *33*(1), 216–234.

Lickteig, T. (1985). Typical tensorial rank. *Linear algebra and its applications*, *69*, 95–120.

Lister, J., Duval, B., Farthing, J., Fredian, T., Greenwald, M., How, J., Llobet, X., Saint-Laurent, F., Spears, W., & Stillerman, J. (2003). The iter project and its data handling requirements, In *9th icalepcs conference, gyeongju, korea*.

Petersen, K. B., Pedersen, M. S. Et al. (2008). The matrix cookbook. *Technical University of Denmark*, *7*(15), 510.

Petty, C. Et al. (2019). Diii-d research towards establishing the scientific basis for future fusion reactors. *Nuclear Fusion*, *59*(11), 112002.

Stegeman, A. (2008). Low-rank approximation of generic $p \times q \times 2$ arrays and diverging components in the candecomp/parafac model. *SIAM Journal on Matrix Analysis and Applications*, *30*(3), 988–1007.

Strassen, V. (1983). Rank and optimal computation of generic tensors. *Linear algebra and its applications*, *52*, 645–685.

Tu, J. H. (2013). *Dynamic mode decomposition: Theory and applications* (Doctoral dissertation). Princeton University.

Udell, M., & Townsend, A. (2019). Why are big data matrices approximately low rank? *SIAM Journal on Mathematics of Data Science*, *1*(1), 144–160.

Vizing, V. G. (1963). The cartesian product of graphs. *Vycisl. Sistemy*, *9*(30-43), 33.

Wang, Z., Lai, M.-J., Lu, Z., Fan, W., Davulcu, H., & Ye, J. (2015). Orthogonal rank-one matrix pursuit for low rank matrix completion. *SIAM Journal on Scientific Computing*, *37*(1), A488–A514.

Zhang, F. (2006). *The schur complement and its applications* (Vol. 4). Springer Science & Business Media.