Low-Rank & Structured Low-Rank Reconstruction Approaches

Image Reconstruction
ISMRRM 2021

Mark Chiew (mark.chiew@ndcn.ox.ac.uk)
Declaration of Financial Interests or Relationships

Speaker Name: Mark Chiew

I have no financial interests or relationships to disclose with regard to the subject matter of this presentation.
Goals

Intuition and Mechanics
A Motivating Example
# The Netflix Problem

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **2**
- **3**
- **4**
- **5**
The Netflix Problem
## The Netflix Problem

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
</tr>
</tbody>
</table>

The table represents user ratings for various movies. Each cell contains a rating from 1 to 5, with 5 being the highest rating. The diagram on the left shows the selection process for choosing movies based on user preferences.
The Netflix Problem

Drama/Horror

Sci-Fi

Action/Comedy
The Netflix Problem
Imaging Example

Users $\rightarrow$ Voxels

Movies $\rightarrow$ Time
Imaging Example

Spatial Component 1  Spatial Component 2  Spatial Component 3  Spatial Component 4

Temporal Component 1  Temporal Component 2  Temporal Component 3  Temporal Component 4
Imaging Example

Spatial Component 1

Spatial Component 2

Spatial Component 3

Spatial Component 4

Temporal Component 1

Temporal Component 2

Temporal Component 3

Temporal Component 4
Imaging Example

Original

Rank-4 Approximation
Background
Singular Value Decomposition

\[ \text{Left Singular Vectors} \times \text{Singular Values} \times \text{Right Singular Vectors} \]
Singular Value Decomposition

\[
\begin{bmatrix}
\end{bmatrix} =
\begin{bmatrix}
U & \Sigma & V^H
\end{bmatrix}
\]

- **Left Singular Vectors**: \( U \)
- **Singular Values**: \( \Sigma \)
- **Right Singular Vectors**: \( V^H \)
Singular Value Decomposition

\[
\text{Left Singular Vectors} \quad \times \quad \Sigma \quad \times \quad \text{Right Singular Vectors}
\]
Singular Value Decomposition

\[ U \times \Sigma \times V^H \]

- **Left Singular Vectors**
- **Singular Values**
- **Right Singular Vectors**
Singular Value Decomposition

\[
U \times \Sigma \times V^H
\]

- **Left Singular Vectors**
- **Singular Values**
- **Right Singular Vectors**
Singular Value Decomposition

Any column of $M$ is a linear combination of columns in $U$.
Singular Value Decomposition

Any row of $M$ is a linear combination of rows in $V^H$
Singular Value Decomposition

Any row of $M$ is a linear combination of rows in $V^H$

Row space basis, defines the row subspace

Linear combination weights
Singular Value Decomposition

Subspaces

Row space basis, defines the row subspace
Singular Value Decomposition

Decomposition into sum of rank-1 matrices
Singular Value Decomposition
Rank-1 matrix

- Singular Value (scalar)
- Right Singular Row Vector
- Left Singular Column Vector
Singular Value Decomposition
Rank-1 matrix
Singular Value Decomposition

Rank-1 matrix

Singular Value (scalar)  Right Singular Row Vector

Left Singular Column Vector
Singular Value Decomposition

Rank-1 matrix
Singular Value Decomposition

Rank-1 matrix
Singular Value Decomposition

Rank-1 matrix

- Singular Value (scalar)
- Right Singular Row Vector
- Left Singular Column Vector
- Left Singular Column Vector (Voxels)
- Right Singular Row Vector (Time)
Singular Value Decomposition

Decomposition into sum of rank-1 matrices

\[
\begin{array}{ccc}
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{bmatrix}
= & \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{bmatrix}
+ & \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{bmatrix}
+ & \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9 \\
\end{bmatrix}
\end{array}
\]
Singular Value Decomposition
Singular Value Decomposition
as a superposition of features

Component 1
Component 2
Component 3
Component 4
Singular Value Decomposition
as a superposition of features

Component 1
Component 2
Component 3
Component 4
Rank, Redundancy and Linear Dependence

The **rank** of a matrix is:

- the number of non-zero singular values
- the number of linearly independent rows and columns in the data
- the dimension of the row and column spaces
Rank, Redundancy and Linear Dependence

- When a matrix is **low-rank**, we can say that:
  - there is information shared across entries of the matrix
  - the rows and columns of the matrix are not linearly independent
  - the rows and columns of the matrix are correlated
  - the rows and column vectors lie in low-dimensional subspaces
  - the matrix is redundant
Matrix Degrees of Freedom
How much information is in a matrix?

8 × 8 matrix
64 DOF
Matrix Degrees of Freedom
How much information is in a matrix?

$$DOF = r(m + n - r)$$

39 DOF
Matrix Degrees of Freedom
How much information is in a matrix?

\[
\begin{bmatrix}
U & \Sigma & V
\end{bmatrix}
\]
Matrix Degrees of Freedom
How much information is in a matrix?

\[ \text{U} \]
\[ \Sigma \]
\[ \text{V} \]

7 + 6
Matrix Degrees of Freedom

How much information is in a matrix?

\[ \begin{array}{ccc}
\text{U} & \text{Σ} & \text{V} \\
\end{array} \]

\[ 7 + 6 + 5 \]
Matrix Degrees of Freedom

How much information is in a matrix?
Matrix Degrees of Freedom

How much information is in a matrix?
Matrix Degrees of Freedom
How much information is in a matrix?

U
∑
V
18
3
7 + 6
Matrix Degrees of Freedom

How much information is in a matrix?

\[ DOF = r(m + n - r) \]

39 DOF
Representational Power

Examples of signals formed by linear combinations of 2 vectors
Connection to Sparsity

$A = U \Sigma V^H$

- **Left Singular Vectors**
- **Singular Values**
- **Right Singular Vectors**
Connection to Sparsity

\[ \text{Left Singular Vectors} \times \text{Singular Values} \times \text{Right Singular Vectors} \]
Connection to Sparsity

\[ \begin{array}{c}
\text{Left Singular Vectors} \\
\text{Singular Values} \\
\text{Right Singular Vectors}
\end{array} \]

Diagonal Vector of Singular Values is Sparse
Connection to Sparsity

$U^H \times \sum \times V = \sum$

$U^H$ and $V$ are like sparsifying transforms

But we don’t need to know them \textit{a priori}, only that they \textit{exist}
Low-Rank Methods
Basic Premise

• You want to reconstruct your data in the presence of under-sampling, noise, or other corruption/perturbation

• Conventional reconstruction models find the image(s) that best fit the measured k-space, in a least squares sense

• However, this may not produce the “best” reconstruction in terms of metrics such as MSE, or the problem might be under-determined

• If your data can be represented as a low-rank matrix, you can use that prior knowledge to additionally constrain/regularize the image reconstruction
Recipe
Low-Rank Methods

1. Identify the low-rank matrix data
   • e.g., space by time

2. Construct your forward measurement model
   • i.e., information about sampling, coil sensitivities, etc.

3. Choose a suitable cost function with low-rank constraint
   • e.g., convex of non-convex, local or global low-rank constraints

4. Solve the cost function using an appropriate algorithm
   • typically a first order, iterative algorithm
1. Identify appropriate matrix data
e.g. (Space $\times$ Time)

- Space-time matrices are probably the most intuitive
- Signal time-courses are correlated across space
- Many voxels exhibit the same contrast dynamics

Liang, “Spatiotemporal imaging with partially separable functions” ISBI 2007
1. Identify appropriate matrix data
e.g. (Space × Spectra)

- Spectral content is similar across voxels

Lam & Liang, “A subspace approach to high-resolution spectroscopic imaging” MRM 2014
1. Identify appropriate matrix data
e.g. (Space × Coils)

- Local spatial information is shared across coils

Trzasko & Manduka, “Calibrationless parallel MRI using CLEAR” ASILOMAR 2011
1. Identify appropriate matrix data
e.g. (Space × TE)

- Voxels or spatial features are correlated across echo time

*Xiaozhi Zhang et al., “Accelerating parameter mapping with a locally low rank constraint” MRM 2014*
1. Identify appropriate matrix data
e.g. (Space $\times$ Diffusion Encoding)

- Voxel contrasts are correlated across diffusion encodings

Gao et al., “PCLR: Phase-constrained low-rank model for compressive diffusion-weighted MRI” MRM 2013
1. Identify appropriate matrix data
n.b. “Casorati Matrix”
1. Identify appropriate matrix data
n.b. “Casorati Matrix”
1. Identify appropriate matrix data

n.b. “Casorati Matrix”
2. Construct the Forward Operator
Maps your “image” to the sampled k-space measurements
2. Construct the Forward Operator
Maps your “image” to the sampled k-space measurements
3. Choose a Low-Rank Constraint
How will the low-rank constraint be enforced?

- Convex Nuclear Norm
  - Sum of singular values (analogous to L1)

- Non-Convex rank constraint
  - Matrix has exactly (or at most) rank “r”

- Global Low-Rank constraint
  - Entire matrix is low-rank

- Locally low-rank constraint
  - Matrix is “patch-wise” low-rank

Convex / Non-Convex
Global / Local
3. Choose a Low-Rank Constraint

Nuclear Norm (Convex)

\[
\|x\|_* = \sum_i \sigma_i
\]

\[
\min_x \|Ex - d\| + \lambda \|x\|_*
\]
3. Choose a Low-Rank Constraint

Strict Rank Constraint (Non-Convex)

\[
\min \quad \|Ex - d\| \\
\text{subject to} \quad \text{rank}(x) \leq r
\]
3. Choose a Low-Rank Constraint

Global Low-Rank Constraint

\[
\min_x \|Ex - d\| + \lambda \|x\|_*
\]
3. Choose a Low-Rank Constraint

Locally Low-Rank Constraint

\[
\min_x \|Ex - d\| + \lambda \sum_p \|R_p x\|_*
\]
3. Choose a Low-Rank Constraint

Extensions with other constraints or formulations

\[
\begin{align*}
\text{Low Rank + Sparse Decomposition}^1 & \quad \min_{L,S} \|E(L + S) - d\| + \lambda_L \|L\|_\ast + \lambda_S \|\Psi S\|_1 \\
\text{Low Rank & Sparse Penalties}^2 & \quad \min_x \|Ex - d\| + \lambda_L \|x\|_\ast + \lambda_S \|\Psi x\|_1 \\
\text{Low Rank Tensor Reconstruction}^3 & \quad \min_x \|Ex - d\| + \lambda \sum_i \|x_{(i)}\|_\ast
\end{align*}
\]

1. Otazo et al., “Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components.” MRM 2015
4. Algorithm
Optimize the cost function with some algorithm

- SVD-free: e.g. Matrix Factorization

- SVD-based: e.g. Proximal gradient method (Singular Value Thresholding)
4. Algorithm
Matrix Factorization - Low Rank by Construction

\[
\min_{U \in \mathbb{C}^{m \times r}} \| E(UV^H) - d \|
\]

\[
V_{k+1} = \arg \min_V \| E(U_{k+1}V^H) - d \|_F
\]

\[
U_{k+1} = \arg \min_U \| E(UV^H_k) - d \|_F
\]

4. Algorithm
Matrix Factorization - Training data

\[
\min_{U \in \mathbb{C}^{m \times r}} \| E(UV^H) - d \| 
\]

- \( V \) can be determined from training data
- e.g. low-resolution, fully-sampled k-space to estimate temporal basis vectors

Liang, “Spatiotemporal imaging with partially separable functions” ISBI 2007
4. Algorithm

SVD-based thresholding

- Analogous to soft-thresholding algorithms for sparse reconstruction problems
- e.g. Alternating between gradient step w.r.t. data consistency and a singular value soft thresholding proximal projection step to promote low-rankness

\[ y_{k+1} = x_k + \mu E^H (E x_k - d) \]
\[ x_{k+1} = \tau_{\lambda \mu} (y_{k+1}) \]

\[ \tau_\alpha(z) = U \cdot \max(\Sigma - \alpha, 0) \cdot V^H \]

Ma et al., “Fixed point and Bregman iterative methods for matrix rank minimization” Math Prog 2011
Other Considerations
Sampling and Parameter Selection

- Sampling
  - With training data (e.g. low-resolution fully-sampled reference), under-sampling can be regular
  - Otherwise, typically more “random” or incoherent sampling, similar to compressed sensing

- Parameter Selection
  - Can be estimated from training data or other reference data, if available
  - Otherwise often empirically chosen, although model-order penalties can also be used

Haldar & Liang, “Spatiotemporal imaging with partially separable functions: A matrix recovery approach” ISBI 2010
Structured Low-Rank Methods
Structured Low Rank Methods

- Similar to general low-rank methods, but now it is not the data matrix itself that is low-rank, but rather a transformation of the input data.

- Don’t require multi-dimensional data, you can generate a structured low-rank matrix from a single image.

- Typically involves collecting points of k-space using a kernel and concatenating them to form a Hankel structured matrix.
Structured Low Rank Methods
The Hankel Matrix

- Take small windows or kernels and unravel them into vectors
Structured Low Rank Methods
The Hankel Matrix

- Take small windows or kernels and unravel them into vectors

\[ \mathcal{H} = \text{Block-Hankel Matrix} \]
Structured Low Rank Methods

The Hankel Matrix

- Stack, or concatenate the vectorized kernels

\[ \mathcal{H}_k = \begin{pmatrix} \text{k-space} \\ \end{pmatrix} \]
Structured Low Rank Methods

The Hankel Matrix

- Stack, or concatenate the vectorized kernels

\[ \mathcal{H} \]
Structured Low Rank Methods
The Hankel Matrix

• Stack, or concatenate the vectorized kernels
Structured Low Rank Methods

The Hankel Matrix

- This procedure generates a block-Hankel structured matrix
Structured Low Rank Methods
The Hankel Matrix

- This procedure generates a block-Hankel structured matrix

\[ \mathcal{H}(\ ) = \text{Block-Hankel Matrix} \]
Structured Low Rank Methods
Convolution Perspective

\[ x \ast y = \mathcal{H}(x) y \]
Structured Low Rank Methods
Limited Support Intuition

Haldar, “Low-Rank Modeling of Local k-Space Neighborhoods (LORAKS) for Constrained MRI” IEEE-TMI 2014
Structured Low Rank Methods
Limited Support Intuition

\[ A \times w - A = 0 \]
\[ A \times (w - 1) = 0 \]
\[ A \times \tilde{w} = 0 \]
\[ \mathcal{F}[A] \otimes \mathcal{F}[\tilde{w}] = 0 \]
\[ \mathcal{H}(\mathcal{F}[A]) \cdot \text{vec}(\mathcal{F}[\tilde{w}]) = 0 \]

Implies the Hankel-structured matrix \( \mathcal{H}(\mathcal{F}[A]) \) has a non-trivial null-space
Which in turn means that \( \mathcal{H}(\mathcal{F}[A]) \) has low-rank
Structured Low-Rank Methods

Other Constraints

- Phase Constrained
  - Haldar, “Low-Rank Modeling of Local k-Space Neighborhoods (LORAKS) for Constrained MRI” IEEE-TMI 2014
  - Lee et al., “Reference-Free Single-Pass EPI Nyquist Ghost Correction Using Annihilating Filter-Based Low Rank Hankel Matrix (ALOHA)” MRM 2016
  - Mani et al., “Multi-Shot Sensitivity-Encoded Diffusion Data Recovery Using Structured Low-Rank Matrix Completion (MUSSELS)” MRM 2017

- Coil Sensitivity Constrained
  - Zhang et al., “Parallel Reconstruction Using Null Operations” MRM 2011
  - Shin et al., “Calibrationless Parallel Imaging Reconstruction Based on Structured Low-Rank Matrix Completion” MRM 2014

- Transform Sparsity Constrained
Structured Low Rank Methods

Reconstruction

\[
\min_x \|Ex - d\| + \lambda \|x\|_* \quad \text{and} \quad \min_{\text{rank}(x) \leq r} \|Ex - d\| \leq r
\]

\[
\min_x \|Ex - d\| + \lambda \|H(\mathcal{F}[x])\|_* \quad \text{and} \quad \min_{\text{rank}(H(\mathcal{F}[x])) \leq r} \|Ex - d\| \leq r
\]

Structured Low-Rank Methods
Sampling Considerations

\[ \mathcal{H} \left( \begin{array}{cccc}
\text{Blue} & \text{Light Blue} & \text{Green} & \text{Yellow} \\
\text{Green} & \text{Light Green} & \text{Orange} & \text{Pink} \\
\text{Yellow} & \text{Orange} & \text{Red} & \text{Pink} \\
\text{Pink} & \text{Red} & \text{Blue} & \text{Green} \\
\end{array} \right) = \right( \begin{array}{cccc}
\text{Green} & \text{Light Green} & \text{Orange} & \text{Pink} \\
\text{Orange} & \text{Light Orange} & \text{Red} & \text{Pink} \\
\text{Pink} & \text{Red} & \text{Blue} & \text{Green} \\
\text{Red} & \text{Blue} & \text{Green} & \text{Yellow} \\
\end{array} \right) \]
Structured Low-Rank Methods
Partial-Fourier Sampling

\[ \mathcal{H}( \cdot ) = \]

\[
\begin{array}{cccccccc}
\hline
& & & & & & & \\
& & & & & & & \\
& & & & & & & \\
\hline
& & & & & & & \\
& & & & & & & \\
& & & & & & & \\
\hline
& & & & & & & \\
& & & & & & & \\
& & & & & & & \\
\end{array}
\]
Structured Low-Rank Methods
Random Contiguous Missing Chunk

\[ H(\cdot) = \]
Structured Low-Rank Methods
Arbitrary Sampling

Haldar, "Low-rank modeling of local k-space neighborhoods: from phase and support constraints to structured sparsity," Proc. SPIE 9597, Wavelets and Sparsity XVI, 2015
Other Resources

• IEEE Signal Processing Magazine, vol 37 no 1, 2020: “Computational MRI”
  • Christodoulou & Lingala, “Accelerated Dynamic Magnetic Resonance Imaging Using Learned Representations: A New Frontier in Biomedical Imaging”
  • Haldar & Setsompop, “Linear Predictability in Magnetic Resonance Imaging Reconstruction: Leveraging Shift-Invariant Fourier Structure for Faster and Better Imaging”