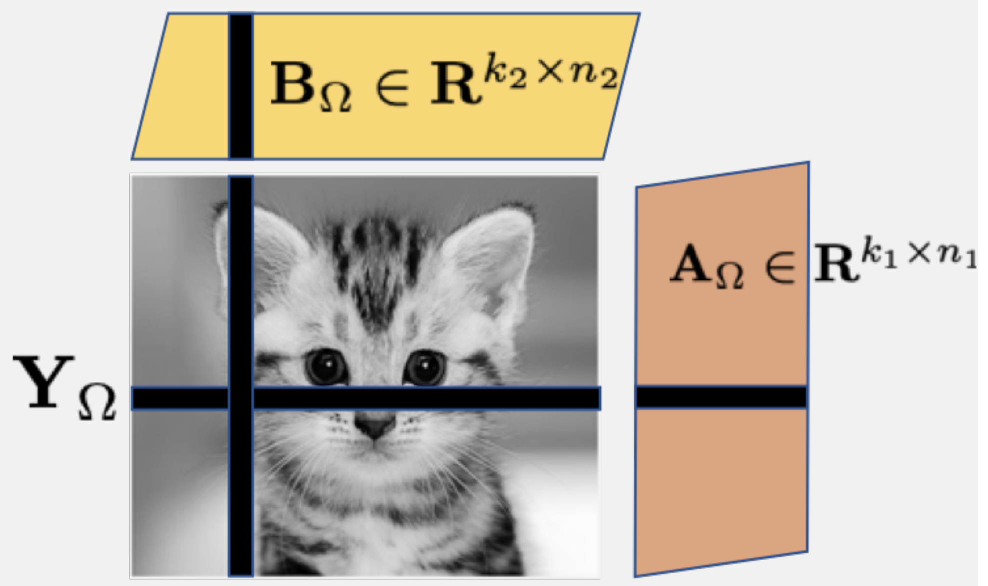
Tensor Matched Kronecker-Structured Subspace Detection for Missing Information

Ishan Jindal and Matthew Nokleby
Wayne State University
{ishan.jindal, Matthew.nokleby}@wayne.edu

Motivation: K-S subspace models for signals with missing information

High Dimensional signals such as human faces can be classified by projecting it on to the low dimensional subspace or union of subspaces. But the signals which are high and **multidimensional** (spatio-Temporal) such as dynamic scene videos, EEG signals or tomographic images have structure in space and time.

- This inherent structure in multidimensional signal can be used to enhance the detection probability of these signals.
- Kronecker-structured (K-S) subspace models are good at exploiting the multidimensional structure.



32x32

This work addresses following questions:

- Information-theoretic analysis of the **detection performance** of K-S models in terms of **residual energy.**
- Matched subspace detection for signal with missing information.
- A more general case of **unstructured missing information** from the signal

K-S Signal Model:

K-S signal model constituents dictionaries for each dimension of the signal and thus accounts for the multi dimensional data structure. For a signal **Y** belongs to one of **L** classes

$Y \in \mathbb{R}^{m_1\times m_2}$: Signal of interest
$A_l \in \mathbb{R}^{m_1\times n_1}$: columns subspace basis
$B_l \in \mathbb{R}^{m_2\times n_2}$: rows subspace basis
$X \in \mathbb{R}^{n_1\times n_2}$: coefficients $\quad Z \in \mathbb{R}^{m_1\times m_2}$: model noise $\quad n_1 < m_1, n_2 < m_2$

$$\mathbf{Y} = \mathbf{A}_l\mathbf{X}\mathbf{B}_l^T + \mathbf{Z}$$

Signal **columns** and **rows** lives an low dimensional subspaces. Vectorized signal model is written as:

$$\mathbf{y} = (\mathbf{B}_l \otimes \mathbf{A}_l)\mathbf{x} + \mathbf{z},$$

$$\mathbf{y} = \mathrm{vec}(Y) \in \mathbb{R}^M \quad \mathbf{z} = \mathrm{vec}(\mathbf{Z}) \in \mathbb{R}^M \quad \mathbf{x} = \mathrm{vec}(X) \in \mathbb{R}^N$$

Signal lives in a low dimensional subspace with a K-S Dictionary. For the Gaussian distributed coefficients and i.i.d noise, the class conditional distribution is

$$p(\mathbf{y}|\mathbf{A}_l, \mathbf{B}_l) = \mathcal{N}(0, (\mathbf{B}_l \otimes \mathbf{A}_l)(\mathbf{B}_l \otimes \mathbf{A}_l)^T + \sigma^2 \cdot \mathbf{I})$$

The misclassification probability by maximum likelihood rule

$$P_e = \frac{1}{L}\sum_{l=1}^{L} \Pr(\hat{l} \neq l | \mathbf{y} \sim p(\mathbf{y}|\mathbf{A}_l, \mathbf{B}_l))$$

For each class K-S dictionaries is defined $\quad \mathbf{D}_l = \mathbf{B}_l \otimes \mathbf{A}_l \in \mathbb{R}^{m_1 m_2 \times n_1 n_2}$

Detection Setup:

Null and the alternate hypothesis $\quad \mathcal{H}_0 : \mathbf{Y} \in \mathbf{D}$ and $\mathcal{H}_1 : \mathbf{Y} \notin \mathbf{D}$

Where, $\quad \mathbf{D} = \mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{m_1 m_2 \times n_1 n_2}$

Test Statistics: $\left\|\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A \mathbf{Y}_\Omega \mathbf{U}_\Omega^B\right\|_F^2 \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \eta.$

Noise free case: $\eta = 0$
Signal with noise: $\eta \neq 0$

To calculate $\quad \mathrm{L.B}(H_1) \leq ||Y_\Omega - \hat{Y}_\Omega||_F^2 \leq \mathrm{U.B}(H_0).$

After simplification we need to bound following four bounds individually,

$$\left\|\mathbf{Y}_\Omega - \hat{\mathbf{Y}}_\Omega\right\|_F^2 \leq ||\mathbf{Y}_\Omega||_F^2 - \left(\underbrace{\left\|(\mathbf{A}_\Omega^T\mathbf{A}_\Omega)^{-1}\right\|_F^2}_{\boxed{\mathrm{II}}} \cdot \underbrace{||\mathbf{A}_\Omega^T \mathbf{Y}_\Omega \mathbf{B}_\Omega||}_{\boxed{\mathrm{III}}} \cdot \underbrace{\left\|(\mathbf{B}_\Omega^T\mathbf{B}_\Omega)^{-1}\right\|_F^2}_{\boxed{\mathrm{IV}}}\right).$$

$\boxed{\mathrm{I}}$ $\quad \boxed{\mathrm{II}}$ $\quad \boxed{\mathrm{III}}$ $\quad \boxed{\mathrm{IV}}$

Theorem: Let the discrete entries from the signal be missing, $\delta > 0$

$$k_1 \geq \tfrac{8}{3} n_1 \mu(\mathbf{A}) \log\left(\tfrac{2n_1}{\delta}\right) \text{ and } k_2 \geq \tfrac{8}{3} n_2 \mu(\mathbf{B}) \log\left(\tfrac{2n_2}{\delta}\right),$$

then with probability at least $1 - 8\delta$,

$$\frac{k_1 m_2 + k_2 m_1 - k_1 k_2}{m_1 m_2}\left((1-\alpha) - \frac{m_1 m_2(\beta+1)^2}{2k_1 k_2(1-\gamma_1)(1-\gamma_2)}\left(\frac{n_1}{m_1}\mu(\mathbf{A}) + \frac{n_2}{m_2}\mu(\mathbf{B})\right)\right)||\mathbf{Y} - \mathbf{U}^A\mathbf{Y}\mathbf{U}^B||_F^2 \leq ||\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A\mathbf{Y}_\Omega\mathbf{U}_\Omega^B||_F^2.$$

$$\left\|\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A\mathbf{Y}_\Omega\mathbf{U}_\Omega^B\right\|_F^2 \leq (1+\alpha)\frac{k_1 m_2 + k_2 m_1 - k_1 k_2}{m_1 m_2}||\mathbf{Y}||_F^2.$$

where:

$$\alpha = \sqrt{\frac{2\mu(\mathbf{Y})^2 k_1 k_2}{(k_1 m_2 + k_2 m_1 - k_1 k_2)^2}\log(\tfrac{1}{\delta})},\ \beta = \sqrt{\frac{4\mu(\mathbf{Y})\log(\frac{1}{\delta})}{\left(\frac{m_1}{k_1} + \frac{m_2}{k_2} - 1\right)\left(\frac{m_2}{n_2}\frac{1}{\mu(\mathbf{B})} + \frac{m_1}{n_2}\frac{1}{\mu(\mathbf{A})}\right)}},\ \gamma_1 = \sqrt{\frac{8n_1\mu(\mathbf{A})}{3k_1}\log\left(\frac{2n_1}{\delta}\right)} \text{ and } \gamma_2 = \sqrt{\frac{8n_2\mu(\mathbf{A})}{3k_2}\log\left(\frac{2n_2}{\delta}\right)}$$

Detection Probability: From the Theorems, for a large enough value of k_1 and k_2 and $\delta > 0$ the probability of detection is more than $1 - 8\delta$.

$$\mathbb{P}\left[\left\|\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A\mathbf{Y}_\Omega\mathbf{U}_\Omega^B\right\|_F^2 > 0 \middle| \mathcal{H}_1\right] \geq 1 - 8\delta$$

$$\mathbb{P}\left[\left\|\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A\mathbf{Y}_\Omega\mathbf{U}_\Omega^B\right\|_F^2 > 0 \middle| \mathcal{H}_0\right] = 0$$

For smaller α, β, γ_1 and γ_2 values, and for incoherent row and column subspace

$\mu(\mathbf{A}) = 1$ and $\mu(\mathbf{B}) = 1$

We write the lower bound in Theorem as follows:

$$\frac{2k(m-k)}{m^2}\left(1 - \frac{nm}{k^2}\right)\left\|\mathbf{Y} - \mathbf{U}^A\mathbf{Y}\mathbf{U}^B\right\|_F^2 \leq \left\|\mathbf{Y}_\Omega - \mathbf{U}_\Omega^A\mathbf{Y}_\Omega\mathbf{U}_\Omega^B\right\|_F^2$$

Here the second term in the bound is always less than zero and leads to very high detection probability.

Lemma I: Using the similar notations, we bound the missing signal energy as

$$(1-\alpha)\left(\frac{k_1 k_2}{m_1 m_2}\right)||\mathbf{Y}||_F^2 \leq ||\mathbf{Y}_\Omega||_F^2 \leq (1+\alpha)\left(\frac{k_1 k_2}{m_1 m_2}\right)||\mathbf{Y}||_F^2.$$

with probability at least $1 - 2\delta$.

Lemma III: With high probability $1 - \delta$ and using same notations, we bound the energy of column subspaces as

$$||\mathbf{A}_\Omega^T\mathbf{Y}_\Omega\mathbf{B}_\Omega||_F^2 \leq (\beta+1)^2\frac{1}{2}\frac{k_1 k_2}{m_1 m_2}\left(\frac{n_1}{m_1}\mu(\mathbf{A}) + \frac{n_2}{m_2}\mu(\mathbf{B})\right)||\mathbf{Y}||_F^2$$

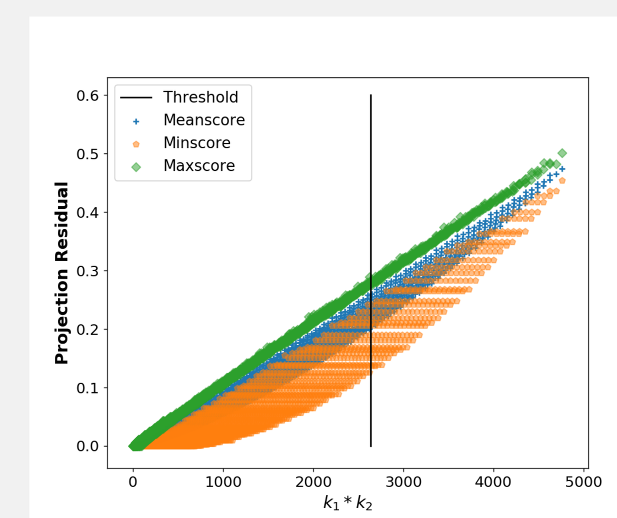
Lemma II and IV: With high probability $1 - \delta$ and using same notation, we bound the energy of column subspaces as

$$\left\|\left(\sum_{i\in\Omega}\mathbf{A}_{\Omega_{i,:}}\mathbf{A}_{\Omega_{i,:}}^T\right)^{-1}\right\|_F \leq \frac{m_1}{(1-\gamma_1)k_1}.$$

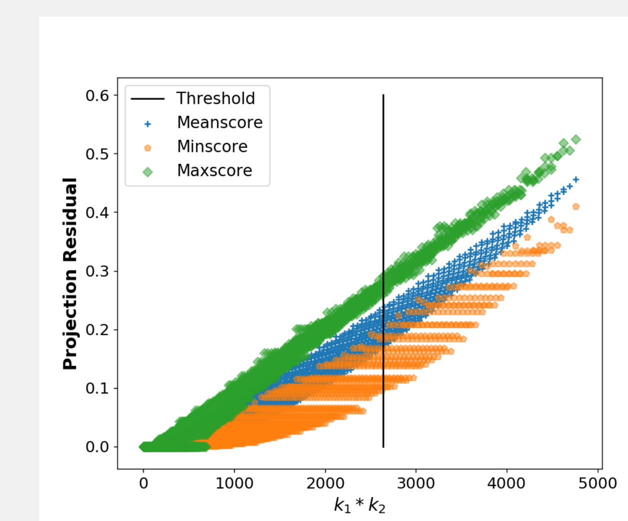
Similarly, we bound the energy of row subspace as

$$\left\|\left(\sum_{i\in\Omega}\mathbf{B}_{\Omega_{:,i}}\mathbf{B}_{\Omega_{:,i}}^T\right)^{-1}\right\|_F \leq \frac{m_2}{(1-\gamma_2)k_2}.$$

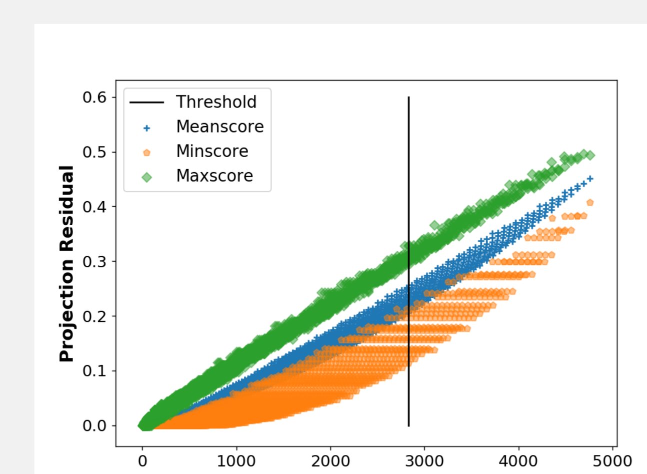
Numerical Results:



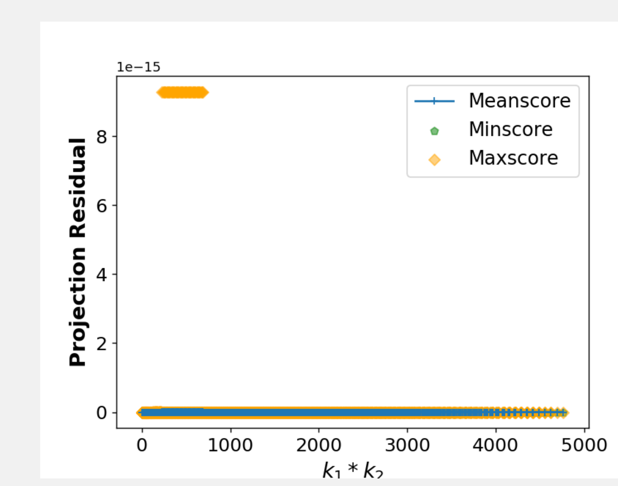
$\mathbf{Y} \in (\mathbf{A} \otimes \mathbf{B})^\perp$



$\mathbf{Y} \in (\mathbf{A}^\perp \otimes \mathbf{B})$



$\mathbf{Y} \in (\mathbf{A} \otimes \mathbf{B}^\perp)$



$\mathbf{Y} \in (\mathbf{A} \otimes \mathbf{B})$

The projection residual $\left\|\mathbf{Y}_{\hat\Omega} - \mathbf{U}_{\hat\Omega}^A\mathbf{Y}_{\hat\Omega}\mathbf{U}_{\hat\Omega}^B\right\|_F^2$ averaged over 1000 simulations for fixed row $\mathbf{A} \in \mathbb{R}^{k_1\times 10}$ and column $\mathbf{B} \in \mathbb{R}^{k_2\times 10}$ subspaces, fixed sample size defined by k_1 and k_2 but different set of samples $\hat\Omega$ drawn without replacement and signal dimension $\mathbf{Y} \in \mathbb{R}^{100\times 100}$.

Conclusion:

We have shown that the detection from the highly incomplete tensor signal is possible by computing the signal energy outside the Kronecker-structured subspace.