Robust Singular Values based on L1-norm PCA

Team: Linear Men

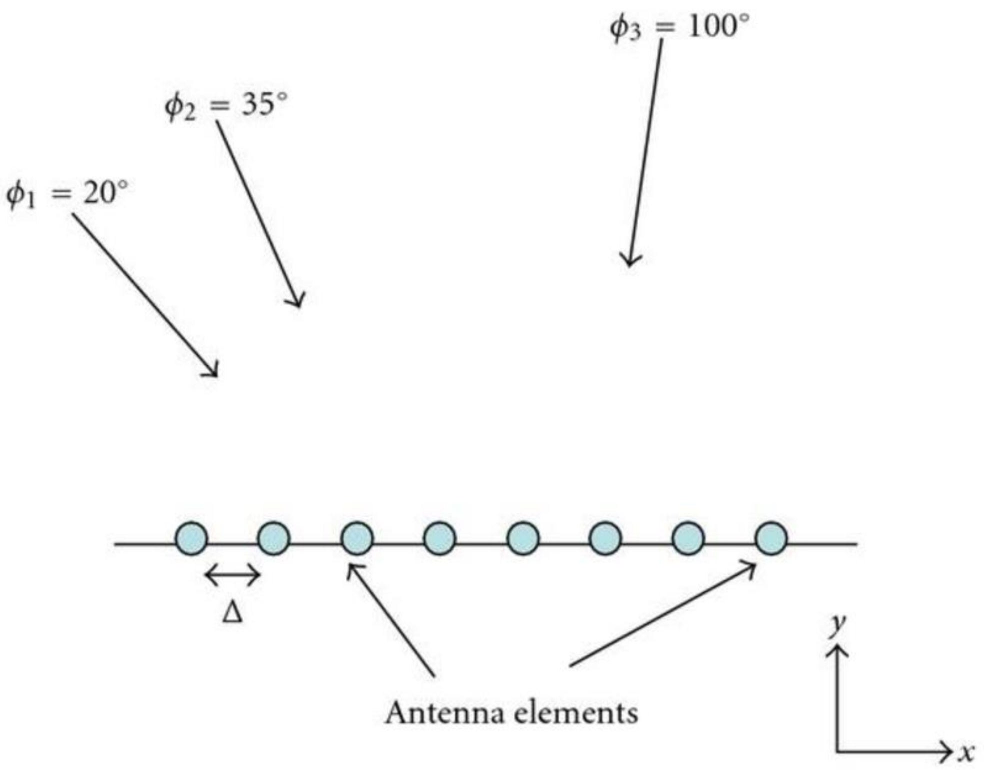
Team members:
- Makhin Artem
- Kuznetsov Mikhail
- Mumladze Maximelian
- Zelentsov Aleksei

Problem

- L2 for PCA is sensitive against outliers
- SVD inherits this outlier sensitivity.
- **We need a robust solution.**
- We implemented SVD and Singular Value (SV) estimation using L1-cSVD.

Robust SV estimation is useful for many engineering problems:

- Channel capacity estimation
- DOA estimation
- Restructuring of Deep NN Acoustic models



Skoltech

Solution

There were solutions for Robust PCA, using L1-PCs
But the results do not diagonalize the data matrix (No SVs)

In this paper, authors utilize previous algos to define a new method called **L1-cSVD** for finding **SVs**

Basically, we want to find $\mathbf{X} \approx \mathbf{U}_{L1}\boldsymbol{\Sigma}_{L1}\mathbf{V}_{L1}^T,$
Where singular vectors are **L1-PCs**
For **X** to be diagonazible, the problem becomes

$$(\boldsymbol{\Sigma}_{L1}, \mathbf{V}_{L1}) = \underset{\substack{\mathbf{V}\in\mathbb{S}^{N\times K},\\ \boldsymbol{\Sigma}\in \mathrm{diag}(\mathbb{R}^K)}}{\mathrm{argmin}} \|\mathbf{X}^T\mathbf{U}_{L1} - \mathbf{V}\boldsymbol{\Sigma}\|_{1,1}.$$

This is a **non-convex opt. problem**, so using different techniques, authors come up with the algorithm on the right

Algorithm 1 L1-cSVD (proposed)

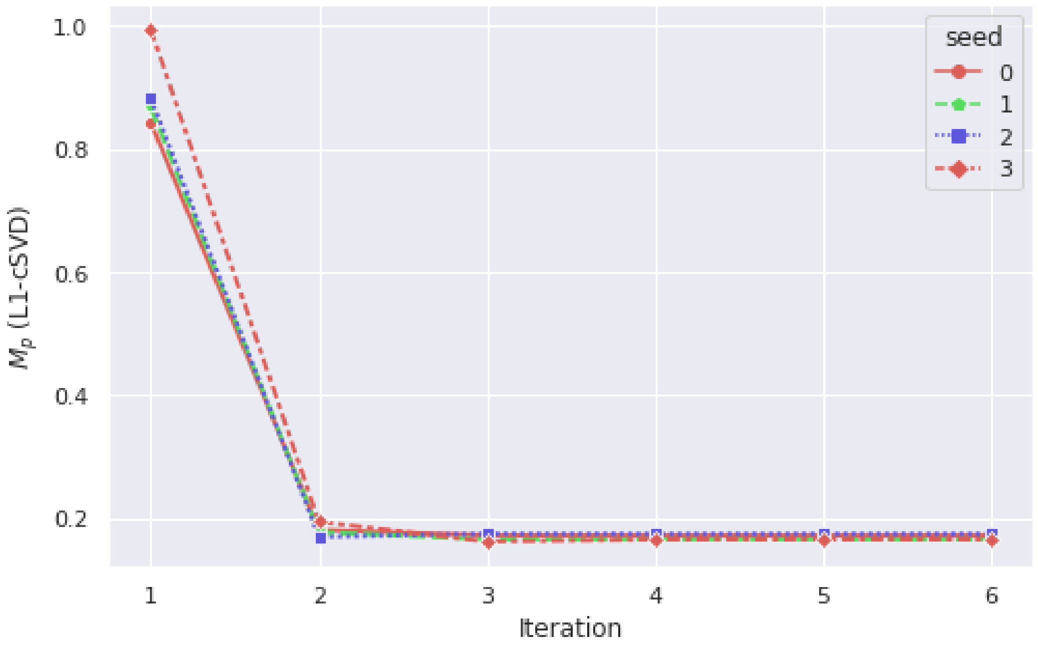
Input: Data matrix $\mathbf{X}_{D\times N}$, number of SVs K
$\mathbf{U} \leftarrow$ L1PCA $(\mathbf{X})$
$\mathbf{A} \leftarrow \mathbf{X}^T\mathbf{U}$
initialization $\boldsymbol{\Sigma} \leftarrow$ zeros(K,K), orthonormal $\mathbf{V}$
while not converged **do**
  for i = 1 to K **do**
    for j = 1 to N **do**
      $\mathrm{s_j} \leftarrow ([\mathbf{A}]_{\mathrm{j,i}}/[\mathbf{V}]_{\mathrm{j,i}})$
      $\mathrm{M_j} \leftarrow \|[\mathbf{A}]_{:,\mathrm{i}} - \mathrm{s}[\mathbf{V}]_{:,\mathrm{i}}\|_1$
    end for
    $\mathrm{j_{opt}} \leftarrow \underset{\mathrm{j}\in[1:N]}{\mathrm{argmin}}\{\mathrm{M_j}\}$
    $[\boldsymbol{\Sigma}]_{\mathrm{i,i}} \leftarrow \mathrm{s_{j_{opt}}}$
  end for
  $(\mathbf{U}', \boldsymbol{\Sigma}', \mathbf{V}') \leftarrow \mathrm{SVD}(\mathbf{A}\boldsymbol{\Sigma}^{-1})$
  $\mathbf{V} \leftarrow \mathbf{U}'\mathbf{V}'^T$
end while
Output: $\mathbf{U}_{L1} \leftarrow \mathbf{U}, \boldsymbol{\Sigma}_{L1} \leftarrow \boldsymbol{\Sigma}, \mathbf{V}_{L1} \leftarrow \mathbf{V}$

Done by Makhin Artem

Skoltech

Experiments, pt.1-2

Convergence Analysis



$$M_P = ||\mathbf{U}^T\mathbf{X} - \mathbf{\Sigma}\mathbf{V}^T||_{1,1} / ||\mathbf{U}^T\mathbf{X}||_{1,1}$$

We see that for all 4 initializations, **L1-cSVD** converges

Done by **Kuznetsov Mikhail**

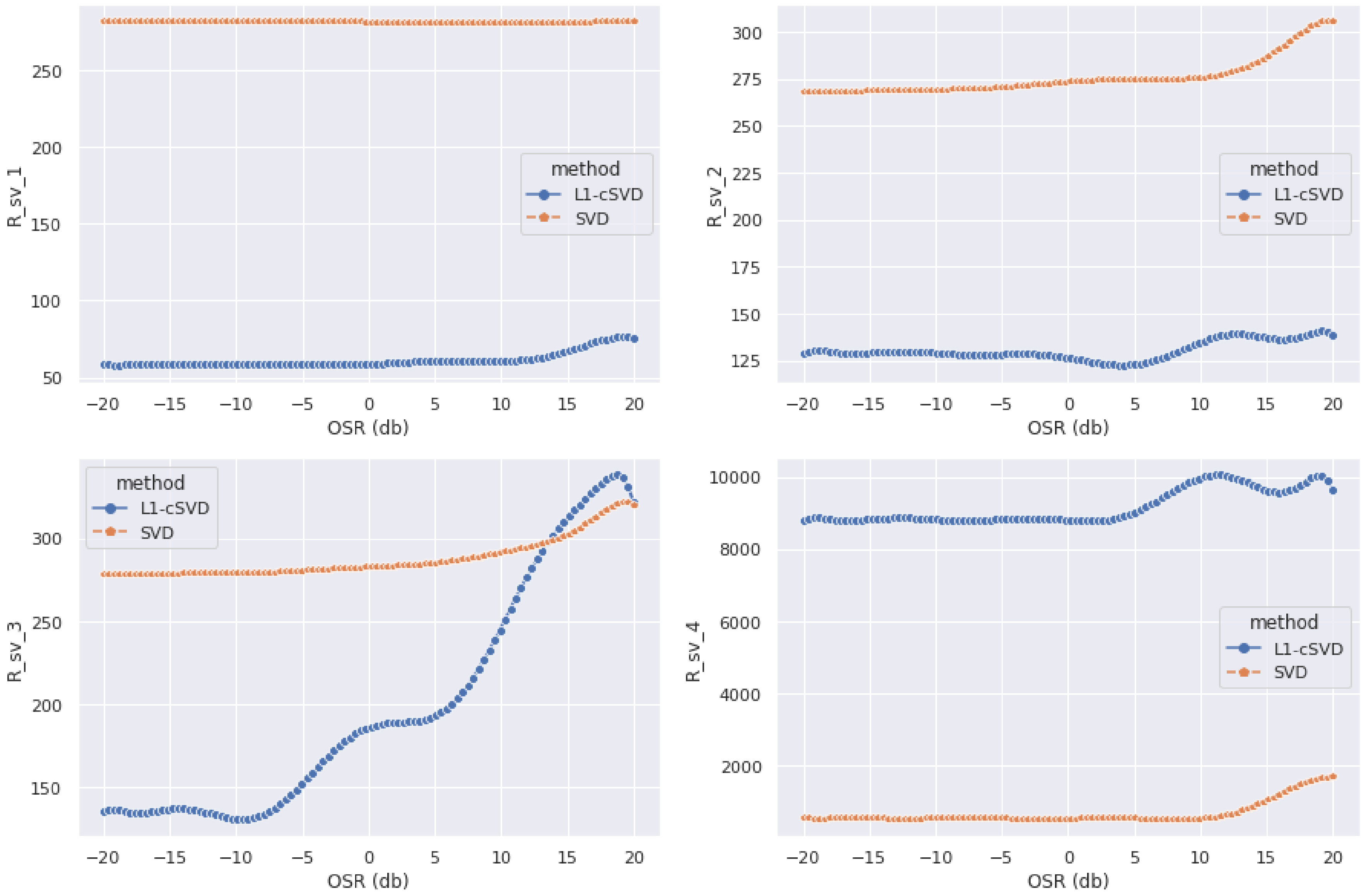
Performance Analysis

We test algos on **SV estimation** using noisy data $R_{\rm sv} = \frac{||\mathbf{\Sigma}^{\rm estimated} - \mathbf{\Sigma}^{\rm clean}||_{2,2}}{||\mathbf{\Sigma}^{\rm clean}||_{2,2}},$

L1-cSVD performs better than **SVD**

Skoltech

Experiments, pt.1-2



Skoltech

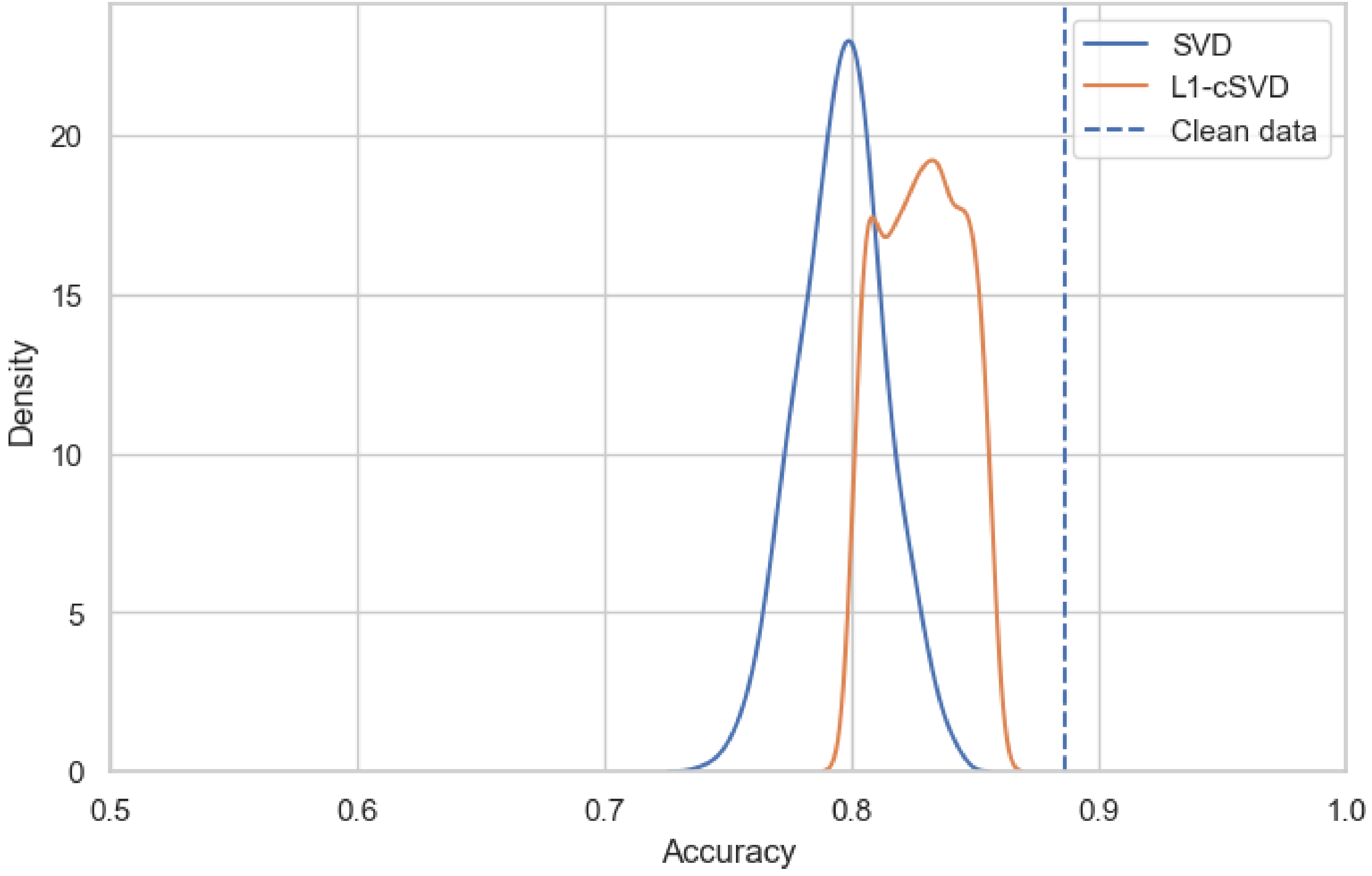
Experiments, pt. 3

Bayesian Classifier

We use Vowel dataset from PMLB, **augmented with outliers**

And then we train a classifier using **standard SVD** and **L1-cSVD**

Our algorithm performs better on corrupted data



Skoltech

Done by Zelentsov Aleksei

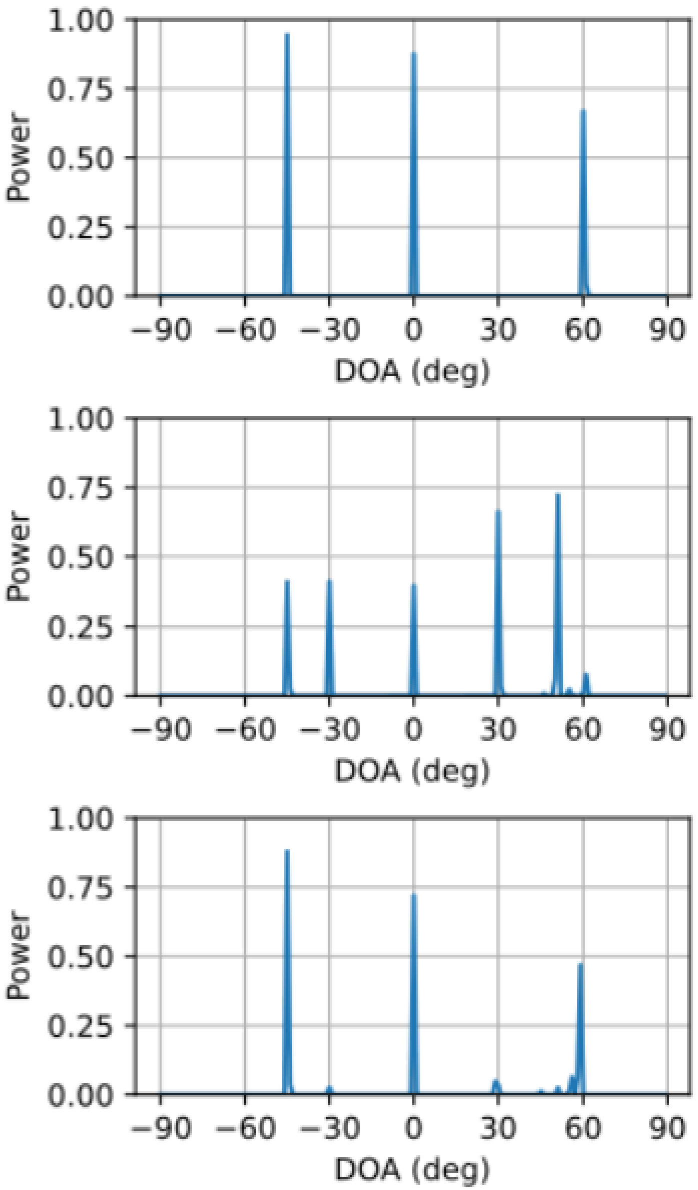
Experiments, pt. 4

Direction of Arrival Estimation

We create incoming signals and add **outlier noise**

We obtain a spatial spectrum through l1-SVD using **standard SVD** and L1-cSVD for dim. reduction

Once again, our algorithm performs better on corrupted data



Skoltech

Done by Mumladze Maximelian

Conclusion

Using several examples, we’ve shown that L1-cSVD performs better Singular Values Estimation on datasets with gross and sparse outliers than other solutions

Skoltech