ECE273 Project Report

Yuchen Zhang[1]

***Abstract*– In this project, I studied the paper *Blind Deconvolution using Convex Programming* with my focus being problem formulation and convex relaxation. I validated the original paper by recreating part of the results. My results show the described deconvolution algorithm are robust to noise and violation of the incoherence assumption. However, breaking time-limitedness of signal w may degrade its performance. Finally, I compared the algorithm to non-blind deconvolution.**

I INTRODUCTION

In the paper *Blind Deconvolution using Convex Programming*[1], the authors aim to separate two signals given their convolution.

In general, the solution to the blind deconvolution problem is non-unique: denote the signal length by L, we have $2L$ unknowns but only L equations. Therefore, the authors added assumptions on the space which the two signals live and proposed the conditions and the algorithm to uniquely recover the two signals.

II PROBLEM FORMULATION

The problem is formulated as follows: Given signals $\mathbf{w}, \mathbf{x}$ of length L and their circular convolution y defined as

$$\mathbf{y} = \mathbf{w} * \mathbf{x}, \quad y[i] = \sum_{j=0}^{L-1} w[j]x[i-j]$$

with the assumptions that

$$\mathbf{w} = \mathbf{B}\mathbf{h}, \qquad \mathbf{h} \in \mathbb{R}^K$$
$$\mathbf{x} = \mathbf{C}\mathbf{m}, \qquad \mathbf{m} \in \mathbb{R}^N$$

where $\mathbf{B} \in \mathbb{R}^{L\times K}$ and $\mathbf{C} \in \mathbb{R}^{L\times N}$. Uniquely recover $\mathbf{w}, \mathbf{x}$.

Here the author made an important assumption: that both convolving signals live in a subspace of higher dimension vector space. Notice that matrix $\mathbf{B}, C$ are not part of recover goal, but merely our tool for representing the dimension of the signal.

III TECHNICAL APPROACH

I preliminaries

I.1 Circulant Matricies

I learned the following from wikipedia [5]:

Circulant matricies can be fully defined by one vector $c \in \mathbb{C}^n$ that have the form

$$\begin{bmatrix} c_0 & c_{n-1} & \cdots & c_1 \\ c_1 & c_0 & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}$$

which can be described as:

$$C_{i,j} = c[(i-j)\%n]$$

Therefore, when multiplying with a vector x:

$$(Cx)_i = \sum_{k=0}^{n-1} C_{i,k} * x[k] = \sum_{k=0}^{n-1} c[(i-k)\%n]x[k]$$

represents a **circular** convolution of vector c and x.

Eigen The wikipedia states the following facts about the eigenvalues and eigenvectors of the circular matrix:

$$v_k = \frac{1}{\sqrt{L}}[1, \omega_0^k, \omega_0^{2k}, \cdots, \omega_0^{(n-1)k}]^T, \quad k = 0, \cdots, n-1$$

Where $\omega_0 = \exp\left(\frac{2\pi i}{n}\right)$ We can express the eigenbasis matrix as:

$$U_{m,n} = \omega_0^{m \times n}$$

With m, n starts from 0 to $n-1$.

Also, the wikipedia claims the eigenvalues are:

$$\begin{aligned}\lambda_k &= c_{n-1}\omega_0^k + \cdots + c_1\omega_0^{(n-1)k} + c_0\underbrace{\omega_0^{nk}}_{=1} \\ &= \sum_{j=0}^{n-1} c_j\omega_0^{(n-j)k} = \sum_{j=0}^{n-1} c_j\omega_0^{-jk}\end{aligned}$$

For $k = 0, \cdots, n-1$

We verify that $Cv_k = \lambda_k v_k$. All index are 0-based.

We first prove the following fact for c, f both periodic with period N.

$$\sum_{j=0}^{N-1} c[(i-j)\%N] * f[j] = \sum_{j=0}^{N-1} c[j] * f[(i-j)\%N]$$

<u>Proof:</u>

$$\begin{aligned}\sum_{j=0}^{N-1} c[(i-j)\%N] * f[j] &= \sum_{j=0}^{N-1} c[(i-j)\%N] * f[j\%N] \\ (l = i-j) \qquad &= \sum_{j=0}^{N-1} c[l\%N] * f[(i-l)\%N] \\ (\text{full period}) \qquad &= \sum_{l=0}^{N-1} c[l\%N] * f[(i-l)\%N]\end{aligned}$$

Expanding all matrix multiplication and use the result above yields:

$$\begin{aligned}(Cv_k)_i &= \sum_{j=0}^{n-1} C_{i,j} * (v_k)_j \\ &= \sum_{j=0}^{n-1} c[(i-j)\%n] * \omega_0^{k \times j} \\ &= \sum_{j=0}^{n-1} c[j] * \omega_0^{(i-j)k} \\ (\lambda_k v_k)_i &= \left(\sum_{j=0}^{n-1} c_j\omega_0^{-jk}\right)\omega_0^{k \times i} \\ &= \sum_{j=0}^{n-1} c[j]\omega_0^{(i-j)k}\end{aligned}$$

Which verifies the validity of eigen vectors and values.

Summarizing the results, in this section we proved that circular convolution of two vectors c, x can be represented as matrix multiplication Cx and obtained a eigen-decomposition of the matrix C.

II Optimization Problem Formulation

II.1 Overview

The author first applies fourier transformation to signals and subspaces involved in the problem to express the operator $circ(\cdot)$ and form a optimization problem from the heuristic about finding the minimal energy signal $\mathbf{h}, \mathbf{m}$ that generate the observation $\mathbf{y}$.

The formed optimization problem is non-convex due to feasible set not a convex set. The author relaxed the problem by taking the dual of primal problem twice. The final result is to not find $\mathbf{h}, \mathbf{m}$ directly, but relax the rank constraint on the input to $\mathcal{A}$, enlarge it from hm^H to all $X \in \mathbb{R}^{K \times N}$ to satisfy the constraint, then minimize the nuclear norm - which by [2] is a heuristic for low rank.

The final result shows the original problem approximated to nuclear norm minimization of matrix X satisfying observation constraint $\mathcal{A}(X) = \hat{\mathbf{y}}$.

II.2 Fourier Transformation

The author first represented the convolution problem as matrix multiplication using circular matricies defined above.

$$\begin{aligned}\mathbf{y} &= m_1(\mathbf{w} * \mathbf{C}_1) + \cdots + m_N(\mathbf{w} * \mathbf{C}_N) \\ &= [circ(\mathbf{C}_1), \cdots, circ(\mathbf{C}_N)][m_1\mathbf{w}, \cdots, m_N\mathbf{w}]^T \\ &= [circ(\mathbf{C}_1)\mathbf{B}, \cdots, circ(\mathbf{C}_N)\mathbf{B}][m_1\mathbf{h}, \cdots, m_N\mathbf{h}]^T\end{aligned}$$

Notice that $circ(\mathbf{C}_i\mathbf{B})$ is a constant that does not depend on the signal variables h, m. Moreover, we can see that the observation is linear w.r.t. the product of the signal variables mh. However, the operator $circ(\cdot)$ is hard to analyze, and we with to express it with known linear algebra terms. Here, fourier decomposition of the circulant matrix comes into play.

From the preliminaries section we can see all circulant matrix have eigenvectors independent of the input vector, and the eigenvalues are linear to the input vector. We can use these elegant properties to represent $circ(\cdot)$ for further analysis.

$$circ(\mathbf{v}) = \mathbf{F}^*\mathbf{\Lambda}\mathbf{F}$$

In which $\mathbf{\Lambda} = diag(\lambda_1, \cdots, \lambda_L)$ and the L-point normalized discrete fourier transform matrix is defined as (0-index)

$$F[\omega, l] = \frac{1}{\sqrt{L}} \exp\left(-2\pi i \frac{\omega l}{L}\right)$$

substitute all $circ(\cdot)$ operator with the above form yields

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{F}\mathbf{y} \\ &= \mathbf{F}[circ(\mathbf{C}_1)\mathbf{B}, \cdots, circ(\mathbf{C}_N)\mathbf{B}][m_1\mathbf{h}, \cdots, m_N\mathbf{h}]^T \\ &= [\mathbf{\Lambda}_1(\mathbf{F}\mathbf{B}), \cdots, \mathbf{\Lambda}_N(\mathbf{F}\mathbf{B})][m_1\mathbf{h}, \cdots, m_N\mathbf{h}]^T \\ &= [\mathbf{\Lambda}_1\hat{\mathbf{B}}, \cdots, \mathbf{\Lambda}_N\hat{\mathbf{B}}][m_1\mathbf{h}, \cdots, m_N\mathbf{h}]^T\end{aligned}$$

In which $\Lambda_i = \sqrt{L}diag(\mathbf{F}C_i) = \sqrt{L}diag(\hat{\mathbf{C}}_i)$

$$\begin{aligned}\lambda_k(circ(\mathbf{C}_i)) &= \sum_{j=0}^{L-1} (\mathbf{C}_i)_j \exp(-jk\frac{2\pi i}{L}) \\ &= \sqrt{L}\sum_{j=0}^{L-1} \mathbf{F}_{k,j} * (C_i)_j \\ &= (\sqrt{L}\mathbf{F}\mathbf{C}_i)_k\end{aligned}$$

Analyzing the l-th element of $\hat{\mathbf{y}}$ from the above matrix multiplication revels:

$$\begin{aligned}\hat{\mathbf{y}}_l &= \sum_{i=1}^{N} (\Lambda_i)_{l,l}\hat{\mathbf{B}}(l,:)m_i\mathbf{h} \\ &= \sum_{i=1}^{N} \sqrt{L}(\hat{\mathbf{C}}_i)_l\hat{\mathbf{B}}(l,:)m_i\mathbf{h} \\ &= \left(\sum_{i=1}^{N} \sqrt{L}(\hat{\mathbf{C}}_i)_l m_i\right)\langle \mathbf{h}, (\hat{\mathbf{B}}^H)_l\rangle \\ &= \langle \underbrace{(\sqrt{L}\hat{\mathbf{C}}^T)_l}_{\hat{\mathbf{c}}_l}, \mathbf{m}\rangle\langle \mathbf{h}, \underbrace{(\hat{\mathbf{B}}^H)_l}_{\hat{\mathbf{b}}_l}\rangle \\ &= \mathrm{tr}(\mathbf{m}^H \underbrace{\hat{\mathbf{c}}_l\hat{\mathbf{b}}_l^H}_{A_l^H}\mathbf{h}) \\ &= \langle \underbrace{\mathbf{h}\mathbf{m}^H}_{X_0}, A_l\rangle\end{aligned}$$

this shows the fourier-transformed observation input can be expressed as a linear combination of the bilinear form of the recover target $\mathbf{h}\mathbf{m}^H$. The author defines this process as a linear operator $\mathcal{A}: \mathbb{R}^{K\times N} \to \mathbb{C}^L$. The problem becomes:

Find matrix X with rank 1 that satisfies $\hat{\mathbf{y}} = \mathcal{A}(X)$.

II.3 Optimization Problem Formulation

We form the problem as finding the vectors h, m with the minimal total energy $\|h\|_2^2 + \|m\|_2^2$ subject to the above equality constraint. This yields the primal optimization problem

$$\min_{h,m} \|h\|_2^2 + \|m\|_2^2 \quad s.t. \quad \hat{y}(l) = \langle hm^H, A_l \rangle, \forall l$$

Although the objective function is a convex function, the feasible set need not be convex as it is a level set of bilinear form of the inputs. Due to the constraint, the problem is not convex.

III Convex Relaxation

III.1 1st Dual

According to paper, the above problem is non-convex due to the feasible set is non-convex. We can convexify it by solving its double Lagrange dual as a approximate of it.

We first take the dual of the primal problem. Notice that we have L **complex** equality constraint, which is in fact $2L$ equality constraints.

Define our multiplier as $a_l, b_l \in \mathbb{R}^L$. For convenience define $d_l = \hat{y}_l - m^H A_l^H h$

$$\begin{aligned}\mathcal{L}((\mathbf{h}, \mathbf{m}), \mathbf{a}, \mathbf{b}) &= \|h\|_2^2 + \|m\|_2^2 \\ &\quad + \sum_{l=1}^{L} (a_l \mathrm{Re}(d_l) + b_l \mathrm{Im}(d_l))\end{aligned}$$

Notice that for convenience we can define $\lambda = \mathbf{a} + \mathbf{b}i$. Therefore we can use operator $\mathrm{Re}(\cdot)$ to simplify our notation:

$$\begin{aligned}\mathrm{Re}(\lambda_i^* d_l) &= \mathrm{Re}\left\{(a_l - b_l i)(\mathrm{Re}(d_l) + \mathrm{Im}(d_l) i)\right\} \\ &= a_l \mathrm{Re}(d_l) + b_l \mathrm{Im}(d_l)\end{aligned}$$

Thus $\mathcal{L}((\mathbf{h}, \mathbf{m}), \mathbf{a}, \mathbf{b})$

$$\begin{aligned}&= \|h\|_2^2 + \|m\|_2^2 + \mathrm{Re}\langle \mathbf{d}, \lambda \rangle \\ &= \|h\|_2^2 + \|m\|_2^2 - \mathrm{Re}\left\{\sum_{l=1}^{L} \lambda_l^* m^H A_l^H h\right\} + \mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle \\ &= \|h\|_2^2 + \|m\|_2^2 - \mathrm{Re}\left\{ m^H \underbrace{\sum_{l=1}^{L} \lambda_l^* A_l^H}_{\tilde{A}^H} h \right\} + \mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle \\ &= \|h\|_2^2 + \|m\|_2^2 - \frac{1}{2}(m^H \tilde{A}^H h + h^H \tilde{A} m) + \mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle \\ &= \begin{bmatrix} h^H & m^H \end{bmatrix} \begin{bmatrix} I_K & -\frac{1}{2}\tilde{A} \\ -\frac{1}{2}\tilde{A}^H & I_N \end{bmatrix} \begin{bmatrix} h \\ m \end{bmatrix} + \mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle\end{aligned}$$

We have the dual problem:

$$\max_{\lambda \in \mathbb{C}^L} \mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle \qquad s.t. \begin{bmatrix} I_K & -\frac{1}{2}\tilde{A} \\ -\frac{1}{2}\tilde{A}^H & I_N \end{bmatrix} \succeq 0$$

Notice that the above is equivalent to (7) in the original paper, as by property of Schur complement [4], as $I_K \succ 0$:

$$\begin{bmatrix} I_K & -\frac{1}{2}\tilde{A} \\ -\frac{1}{2}\tilde{A}^H & I_N \end{bmatrix} \succeq 0 \Leftrightarrow I_N - \frac{1}{4}\tilde{A}^H A \succeq 0 \Leftrightarrow \begin{bmatrix} I_K & \frac{1}{2}\tilde{A} \\ \frac{1}{2}\tilde{A}^H & I_N \end{bmatrix} \succeq 0$$

Scaling entry of λ by half yields:

$$\max_{\lambda \in \mathbb{C}^L} 2\mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle \qquad s.t. \begin{bmatrix} I_K & \tilde{A} \\ \tilde{A}^H & I_N \end{bmatrix} \succeq 0$$

We take the dual of this dual problem to yield a approximate of the primal.

III.2 2nd Dual

We wish to apply result (2.8) in the reference[3]. As the current linear maps are in complex field, while (2.8) states the result about the real field, we need to real-ify the linear maps defined. The below proof to apply the result seems more of a acrobatics in expressing complex operations in real numbers and are quite boring, I leave it in the appendix in section VI.

By work in [3] on the duality between operator norm and trace, the second dual of the primal is:

$$\min_X \|X\|_* \quad s.t. \quad \hat{\mathbf{y}} = \mathcal{A}(X)$$

As the objective function is convex and equality constraint is linear in X, the problem is now convex.

IV Theoretical Guarantees

IV.1 Overview

In the paper, the author gives theoretical guarantee of recovery process of uniquely recovering the rank-1 matrix $X_0 = hm^H$.

The guarantee depends on assumptions on the signal subspace B and signal w in the fourier domain.

To measure effectiveness, the author set the unknown subspace C following isotropic distribution, and gives a probabilistic guarantee of uniquely recover X_0. Under noise, the author also shows guarantee of recover X_0 with noise with energy limited to scalar of the energy of noise in the measurement with the same probability.

IV.2 Assumptions

Specifically, the author made the following assumptions to set up guarantee:

1. WLOG assumptions.

   The author assumed columns of B are orthonormal. and $\|\mathbf{h}\| = 1$.

2. Time-limitedness of signal w

   The author assumed the signal w is time-limited to Q where $K \leq Q \leq L$. This requires the last $L - Q$ rows of B are zero.

3. Incoherence of basis B in fourier domain

   The author states their guarantee is strong when energy is evenly distributed in the rows of $\hat{\mathbf{B}} = \mathbf{FB}$. The extent of incoherence is measured by quantity

   $$\mu^2_{max} = \frac{L}{K} \max_{1 \leq l \leq L} \|\hat{\mathbf{b}}_l\|_2^2$$
   $$\mu^2_{min} = \frac{L}{K} \min_{1 \leq l \leq L} \|\hat{\mathbf{b}}_l\|_2^2$$

   with $\mu^2_{max} = 1$ showing maximum incoherence.

4. Diffusion of signal w in fourier domain

   The author measures diffusion of the signal of interest w by

   $$\mu^2_h = L \max_{1 \leq l \leq L} |\langle \mathbf{h}, \hat{\mathbf{b}}_l \rangle|^2$$

5. Independence of columns of C in fourier domain

   columns of $\hat{\mathbf{C}}$ are independent.

IV.3 Guarantee

Under a generic subspace $\mathbf{C}$ satisfying isotropic normal distribution

$$C_{l,n} \sim \mathcal{N}(0, L^{-1})$$

The author shows that for $\alpha \geq 1$, if:

$$\begin{cases} Q & \geq C_\alpha \cdot M \log(L) \log(M) \\ \frac{L}{Q} & \geq \log_2(C'_\alpha \sqrt{N \log L}) \end{cases}$$

Where $M = \max(\mu^2_{max} K, \mu^2_h N)$Then there exists C''_α that depends only on α such that if

$$M \leq \frac{L}{C''_\alpha \log^3 L}$$

then $X_0 = hm^*$ is the unique solution to the nuclear norm minimization approximation with probability $1 - \mathcal{O}(L^{-\alpha+1})$.

The authors also proved the stability of their method when the measurement y is affected by a energy limiting noise:

Assuming all conditions from the noise-free guarantee, let $\hat{\mathbf{y}} = \mathcal{A}(X_0) + z$ where $z \in \mathbb{R}^L$ is a unknown noise vector with $\|z\|_2 \leq \delta$, then solving the convex program

$$\min \|X\|_* \quad s.t. \quad \|\hat{\mathbf{y}} - \mathcal{A}(X)\|_2 \leq \delta$$

with the same probability $1 - \mathcal{O}(L^{-\alpha+1})$ we can recover $\tilde{X}$ with

$$\|\tilde{X} - X_0\|_F \leq C\frac{\lambda_{max}}{\lambda_{min}}\sqrt{\min(K,N)}\delta$$

i.e., the correct matrix plus a energy limited noise linear with the input noise.

IV RESULTS

I Result Recreation and Analysis

I developed program in python to solve the deconvolution problem by solving

$$\min_X \|X\|_* \quad s.t. \quad \hat{y} = \mathcal{A}(X)$$

using the cvxpy package.

I.1 Phase dirgrams

Limited by computation resource, I scaled down the dimensions to $L = 40$, and N, K varies between 0 to $L/2$.

To generate the signals for the below diagram, I first constructed the basis B, C according to the paper: sparse refers to its columns are randomly chosen from the identity matrix; short refers to its column are the first K or N columns of the identity matrix; generic refers to its element follows iid standard gaussian distribution.

With the constructed matrix, I constructed vector h, m by sampling from iid gaussian distribution. Then I computed $w = Bh$ and $x = Cm$ as the testing signal, and y from w, x.

For each figure, I varied K, N from 1 to 20, repeated each dimension combination for 100 times and recorded the success rate. Here are the results:

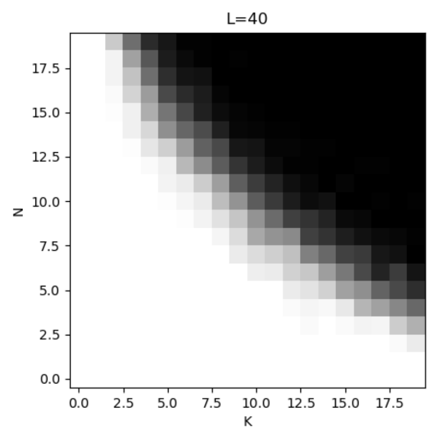

Figure 1: sparse w, generic x

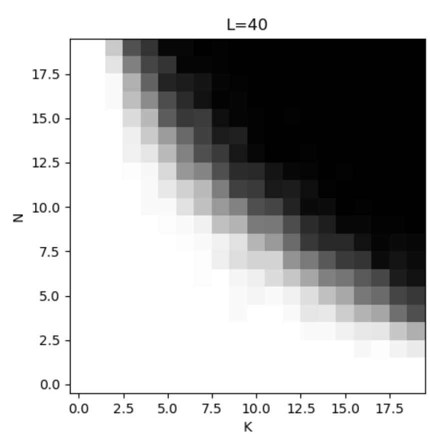

Figure 2: short w, generic x

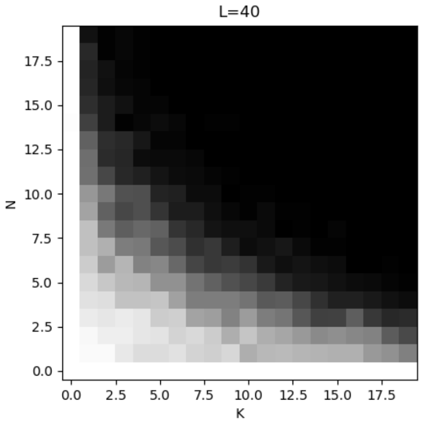

Figure 3: short w, sparse x

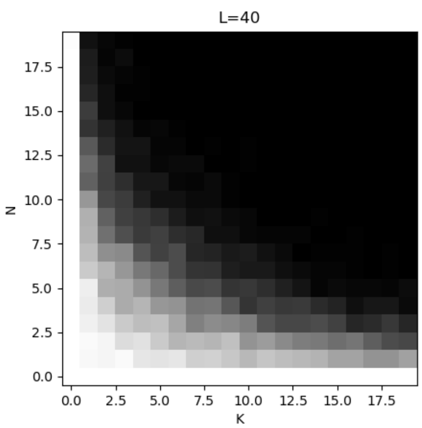

Figure 4: sparse w, sparse x

The figures for generic x accords with the paper, but there are difference in the case where x is sparse, for which the success rate drops significantly. I suspect not enlarging the observation scale may be the issue, but I didn't test my hypothesis due to long computation time.

I.2 Stability

I conducted the stability test as follows: I fixed the dimensions $L = 80$, $N = 20$, $K = 10$. After generating a sparse w and generic x, and its true observation y, I pertubated $\hat{y} = y + z$ where $z \sim N(0, \sigma^2 I)$. I varied the value of σ, computed the input signal-to-noise ratio as $\|y\|_2 / \|z\|_2$, and recorded the relative error of recovered $\hat{X}$ to the ground truth hm^T. Each data point is the averaged result of 100 tests.

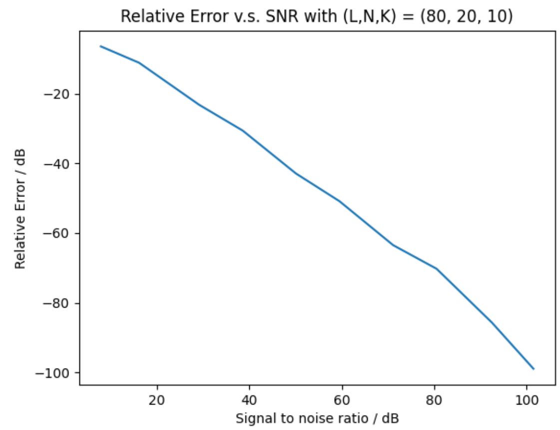

Figure 5: stability of solution under noise

The above result accords with the original paper, showing stability of the recovery. More, we observe a linear relation on the log-log plot with slope approximately -1. This shows the relative noise is approximately in inverse relation with the input signal-noise ratio.

This observation aligned with the author's proof about the stability theorem that the error in $\|X\|_F$ is less then constant times δ.

I then recreated the oversampling experiment. I fixed $N = 20, K = 10$ in this experiment and varied L. I adjusted the value of σ such that the input signal-to-noise ratio is about $20dB$. Each data point is the averaged(averaged in dB and converted bask) result of 100 tests.

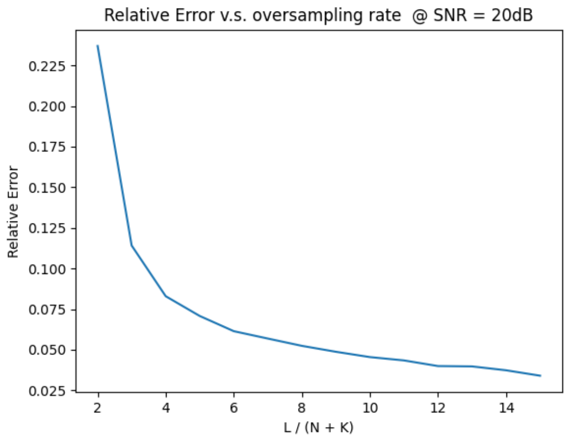

Figure 6: sparse w, sparse x

The result shows relative error decreases with increasing L.

II Comparision to non-blind deconvolution

In this section, we compare the performance difference with/without knowledge of the signal w. With the knowledge of w, the primal problem becomes convex in the constraint and can be directly solved efficiently. In the below experiments, we solved the unknown m by solving:

$$\min_m \|m\|_2 \quad s.t. \quad \hat{y} = \mathcal{A}(hm^*)$$

and evaluate the relative error using the same method as before:

$$\|h(\hat{m} - m)^T\|_F / \|h\hat{m}^T\|_F$$

II.1 probability of success

The phase diagram is trivial here as we have the same number of unknowns as the constraints given signal w. A solution is always guaranteed.

II.2 Stability

I repeated the stability experiments as in the blind deconvolution case with the same parameters. The only difference is that in the presence of noise, the optimization problem may not always be feasible with the δ as before. My solution to this problem is to double δ and retry if problem is infeasible.

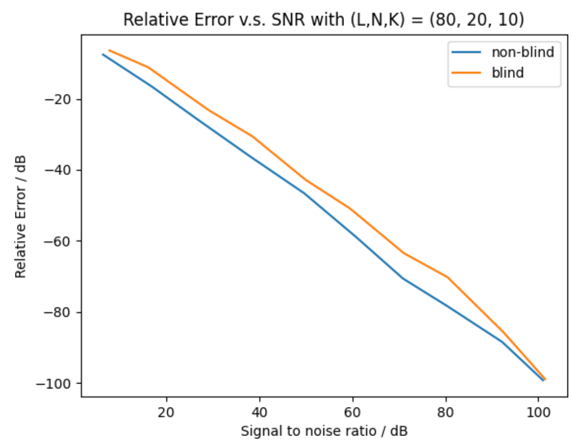

Figure 7: sparse w, sparse x

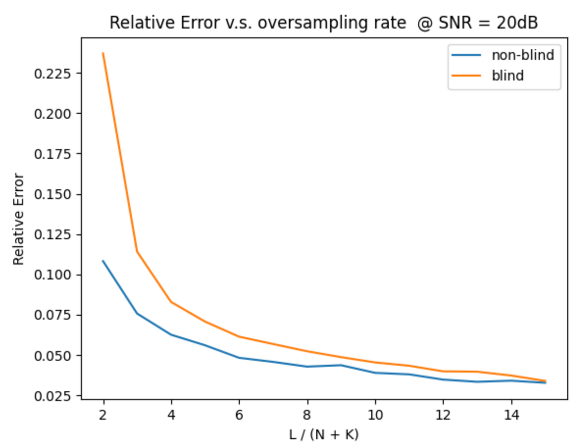

Figure 8: sparse w, sparse x

As expected, the non-blind deconvolution yields a lower relative error on all dimensions. From the first figure we can see the non-blind deconvolution result is also inversely proportionally to input signal-to-noise ratio.

III Evaluation of theoretical guarantees

This section investigates the robustness of the algorithm with respect to the conditions required by the theoretical guarantee section.

In the below experiments, I will generate data of fixed size $L = 40, N = 10, K = 10$ and record the probability of successful deconvoluton of 100 repeated tries.

III.1 Incoherence of $\hat{B}$

I generated B of different ranges of μ_{max} and evaluated their success rate on a generic C matrix with dimensions $L = 40$, $N = 10$, $K = 10$. I split the incoherence range linearly into 20 bins, drew 20 matrix B from the bin, and repeat tested each for 20 times.

I generated matrix B of different incoherence by combining 3 types of basis with random ratio, then normalize the energy. The three types are:

1. Large incoherence: random columns of the identity matrix
2. Medium incoherence: matrix with element sampled from standard gaussian
3. Small incoherence: matrix filled with 1.

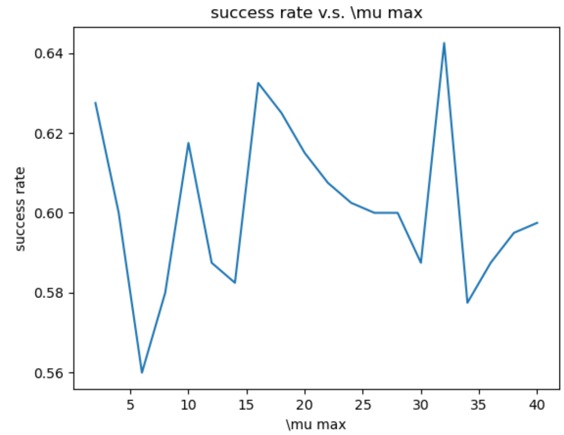

Figure 9: effect of incoherence

The result shows that success rate of the recovery is not influenced significantly by the incoherence of B.

III.2 Time-limitedness of w

To investigate the effect of time-limitedness, I modified the experiment for figure2. Instead of setting entries after K as 0, I set them to a random value, breaking time-limitedness of w. The comparision between phase diagrams are:

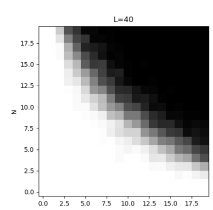

Figure 10: time limited

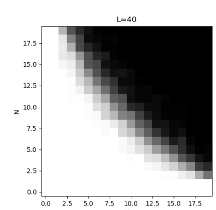

Figure 11: not time limited

We can see the phase diagram of non-time-limited signal is slightly worse than the time limited version. Time-limitedness, by itself, can degrade the performance of the algorithm.

V Conclusion

In conclusion, in this project I studied the paper *Blind Deconvolution using Convex Programming* and re-created some of its results on a smaller scale. Overall, my reproduction results accords with the original paper. Comparision with non-blind deconvolution shows blind version produces more noise and are less sample efficient. Tests on violating theoretical assumptions shows that the algorithm is robust w.r.t. incoherence of B, while breaking time-limitedness can degrade its performance.

For myself, I learned that circulant convolution have a connection to discrete fourier transform via eigen-decomposition, and how the initial problem is relaxed twice into its "convex approximation" and efficiently solved.

References

[1] Ali Ahmed, Benjamin Recht, and Justin Romberg. *Blind Deconvolution using Convex Programming*. 2012. eprint: `arXiv:1211.5608`.

[2] Maryam Fazel. PhD thesis. 2002. URL: `https://faculty.washington.edu/mfazel/thesis-final.pdf`.

[3] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*. June 2007. URL: `https://arxiv.org/abs/0706.4138`.

[4] *Schur complement*. Sept. 2021. URL: `https://en.wikipedia.org/wiki/Schur_complement`.

[5] Wikipedia. *Circulant matrix*. 2022. URL: `https://en.wikipedia.org/wiki/Circulant_matrix` (visited on 05/24/2022).

VI APPENDIX

I Proof of second dual

Separating real and imaginary components,

$$2\mathrm{Re}\langle \hat{\mathbf{y}}, \lambda \rangle = 2 \begin{bmatrix} \mathrm{Re}(\hat{\mathbf{y}})^T & \mathrm{Im}(\hat{\mathbf{y}})^T \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\lambda) \\ \mathrm{Im}(\lambda) \end{bmatrix}$$

Define linear map $\mathcal{F} : \mathbb{R}^{(2K)\times N} \to \mathbb{R}^{2L}$

$$\begin{bmatrix} \mathrm{Re}(\hat{\mathbf{y}}) \\ \mathrm{Im}(\hat{\mathbf{y}}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(\mathcal{A}(X_r + iX_i)) \\ \mathrm{Im}(\mathcal{A}(X_r + iX_i)) \end{bmatrix} = \mathcal{F}\left(\begin{bmatrix} X_r \\ X_i \end{bmatrix} \right)$$

where $X_r, X_i \in \mathbb{R}^{K\times N}$
We wish to show that $\mathcal{V} : \mathbb{R}^{2L} \to \mathbb{R}^{K\times N} \times \mathbb{R}^{K\times N}$ defined as

$$\mathcal{V}\left(\begin{bmatrix} \mathrm{Re}(\lambda) \\ \mathrm{Im}(\lambda) \end{bmatrix} \right) = \begin{bmatrix} \mathrm{Re}(\tilde{A}) \\ \mathrm{Im}(\tilde{A}) \end{bmatrix}$$

is the adjoint linear operator of $\mathcal{F}$.
<u>Proof:</u>

$$\begin{aligned}
\left\langle \begin{bmatrix} v_r \\ v_i \end{bmatrix}, \mathcal{F}\left(\begin{bmatrix} X_r \\ X_i \end{bmatrix} \right) \right\rangle &= v_r^T \mathrm{Re}(\mathcal{A}(X_r + iX_i)) + v_i^T \mathrm{Im}(\mathcal{A}(X_r + iX_i)) \\
&= \mathrm{Re}\, \langle \mathcal{A}(X_r + iX_i), v_r + iv_i \rangle \\
\left\langle \mathcal{V}\left(\begin{bmatrix} v_r \\ v_i \end{bmatrix} \right), \begin{bmatrix} X_r \\ X_i \end{bmatrix} \right\rangle &= tr\left(\begin{bmatrix} \mathrm{Re}(\tilde{A})^T & \mathrm{Im}(\tilde{A})^T \end{bmatrix} \begin{bmatrix} X_r \\ X_i \end{bmatrix} \right) \\
&= tr(\mathrm{Re}(\tilde{A}(v_r, v_i))^T X_r) + tr(\mathrm{Im}(\tilde{A}(v_r, v_i))^T X_i) \\
&= \mathrm{Re}\left(tr(\tilde{A}(v_r, v_i)^H X_r) \right) - \mathrm{Im}\left(tr(\tilde{A}(v_r, v_i)^H X_i) \right) \\
&= \mathrm{Re}\left(tr(\tilde{A}(v_r, v_i)^H X_r) \right) + \mathrm{Re}\left(tr(\tilde{A}(v_r, v_i)^H iX_i) \right) \\
&= \mathrm{Re}\left\{ tr(\tilde{A}(v_r, v_i)^H (X_r + iX_i)) \right\} \\
&= \mathrm{Re}\left\{ \sum_{l=1}^{L} (v_r + iv_i)_l^H tr\left[A_l (X_r + iX_i) \right] \right\} \\
&= \mathrm{Re}\left\{ \sum_{l=1}^{L} (v_r + iv_i)_l^H \mathcal{A}(X_r + iX_i)_l \right\} \\
&= \mathrm{Re}\langle \mathcal{A}(X_r + iX_i), v_r + iv_i \rangle \\
\Rightarrow \mathcal{V} &= \mathcal{F}^*
\end{aligned}$$

Finally we have $\|\mathcal{V}\| = \|\tilde{A}\|$ As

$$\|\mathcal{V}(\lambda)\| = tr(\mathrm{Re}(\tilde{A})^T\mathrm{Re}(\tilde{A}) + \mathrm{Im}(\tilde{A})^T\mathrm{Im}(\tilde{A})) = tr(\tilde{A}^H\tilde{A}) = \|\tilde{A}\|$$

when both operators are given λ as input. Strictly speaking, there exists a norm-preserving bijection between the inputs of both sides.