Phaseless PCA: Low-Rank Matrix Recovery from Column-wise Phaseless Measurements

Seyedehsara Nayer, Praneeth Narayanamurthy, Namrata Vaswani

Iowa State University

Phase Retrieval (PR)

- Recover a length n signal $\boldsymbol{x}^*$ from its phaseless linear projections

$$\boldsymbol{y}_i := |\langle \boldsymbol{a}_i, \boldsymbol{x}^* \rangle|, \quad i = 1, 2, \ldots, m$$

- Without any structural assumptions, PR necessarily needs $m \geq n$.

To reduce sample complexity, can try to exploit structure

- Most existing work studies sparse PR – assumes $\boldsymbol{x}^*$ is sparse.

Phase Retrieval (PR)

- Recover a length n signal $\boldsymbol{x}^*$ from its phaseless linear projections

$$\boldsymbol{y}_i := |\langle \boldsymbol{a}_i, \boldsymbol{x}^* \rangle|, \quad i = 1, 2, \ldots, m$$

- Without any structural assumptions, PR necessarily needs $m \geq n$.

To reduce sample complexity, can try to exploit structure

- Most existing work studies sparse PR – assumes $\boldsymbol{x}^*$ is sparse.
- Another simple structure is low-rank. Two ways to use this:
  1. assume $\boldsymbol{x}^*$ can be rearranged as a low-rank matrix (not studied); or
  2. assume a set of signals (or vectorized images) $\boldsymbol{x}_k^*$, $k = 1, 2, \ldots, q$, together form a low-rank matrix

The second is a more practical and commonly used model and we use this:

- first studied in our earlier work [Vaswani, Nayer, Eldar, Low-Rank Phase Retrieval, T-SP'17]

Recover an $n \times q$ matrix of rank r

$$\boldsymbol{X}^* = [\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \dots, \boldsymbol{x}_k^*, \dots \boldsymbol{x}_q^*]$$

from a set of m phaseless linear projections of *each* of its q columns

$$\boldsymbol{y}_{ik} := |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle|, \;\; i = 1, \dots, m, \;\; k = 1, \dots, q.$$

Application: fast phaseless dynamic imaging, e.g., Fourier ptychographic imaging of live biological specimens

Recover an $n \times q$ matrix of rank r

$$\boldsymbol{X}^* = [\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \ldots, \boldsymbol{x}_k^*, \ldots \boldsymbol{x}_q^*]$$

from a set of m phaseless linear projections of *each* of its q columns

$$\boldsymbol{y}_{ik} := |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle|, \ \ i = 1, \ldots, m, \ \ k = 1, \ldots, q.$$

Application: fast phaseless dynamic imaging, e.g., Fourier ptychographic imaging of live biological specimens

- Even the linear version of this problem is different from both
  - LR matrix sensing: recover $\boldsymbol{X}^*$ from $\boldsymbol{y}_i = \langle \boldsymbol{A}_i, \boldsymbol{X}^* \rangle$ with $\boldsymbol{A}_i$'s dense
    - ★ global measurements (y_i depends on entire $\boldsymbol{X}^*$)
  - LR matrix completion: recover $\boldsymbol{X}^*$ from a subset of its entries
    - ★ completely local measurements
    - ★ need rows & cols to be “dense” to allow for correct “interpolation”
- Our problem - non-global measurements of $\boldsymbol{X}^*$, but global for each column
  - only need denseness of rows (incoherence of right singular vectors)

Recover an $n \times q$ rank-r matrix $\boldsymbol{X}^*$ from $\boldsymbol{y}_{ik} = |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle|$, $i \in [1, m]$, $k \in [1, q]$.

AltMinLowRaP algo: careful spectral init followed by alternating minimization.

Theorem (Guarantee for AltMinLowRaP)

Assume μ-incoherence of right singular vectors of $\boldsymbol{X}^*$. Set $T := C \log(1/\epsilon)$. Assume that, for each new update step, we use a new (independent) set of mq measurements with m satisfying

$$mq \geq C\kappa^6 \mu^2 \ nr^4$$

and $m \geq C \max(r, \log q, \log n)$. Then, w.p. at least $1 - 10n^{-10}$,

$$dist(\hat{\boldsymbol{x}}_k^T, \boldsymbol{x}_k^*) \leq \epsilon \|\boldsymbol{x}_k^*\|, \ k = 1, 2, \ldots, q$$

Also, the error decays geometrically with t.

Sample complexity: $C \cdot nr^4 \log(1/\epsilon)$ (treating κ, μ as constants).
Time complexity: $C \cdot mqnr \log^2(1/\epsilon)$.

Recover a rank-r $n \times q$ matrix $\boldsymbol{X}^*$ from $\boldsymbol{y}_{ik} = |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle|$, $i \in [1, m]$, $k \in [1, q]$.

- Treating κ, μ as constants, our sample complexity is

$$m_{\text{tot}} q \geq C \; n r^4 \log(1/\epsilon)$$

  Number of unknowns in $\boldsymbol{X}^*$ is $(q + n) r \approx 2nr$

  - sample complexity is r^3 times the optimal value (nr)

- No existing guarantees for our problem or even its linear version:
  - closest LR recovery problem with non-global measurements is LR Matrix Completion (LRMC)

Recover a rank-r $n \times q$ matrix $\boldsymbol{X}^*$ from $\boldsymbol{y}_{ik} = |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle|$, $i \in [1, m]$, $k \in [1, q]$.

- Treating κ, μ as constants, our sample complexity is

$$m_{\text{tot}} q \geq C \; n r^4 \log(1/\epsilon)$$

  Number of unknowns in $\boldsymbol{X}^*$ is $(q+n)r \approx 2nr$
  - sample complexity is r^3 times the optimal value (nr)
- No existing guarantees for our problem or even its linear version:
  - closest LR recovery problem with non-global measurements is LR Matrix Completion (LRMC)
- Sample complexity of non-convex LRMC solutions is also sub-optimal
  - AltMinComplete needs $C \; n r^{4.5} \log(1/\epsilon)$ samples
  - Best LRMC solution (proj-GD) needs $C \; n r^2 \log^2 n$ samples
- Comparison with standard (unstructured) PR
  - Standad PR sample complexity is nq: much larger when $r^4 \ll q$

Key idea of the algorithm: Alt-Min for Phaseless Low Rank Recovery [Nayer, Narayanamurthy, Vaswani, Phaseless PCA, ICML 2019 (this work)]

- Alternating minimization relies on the following key idea:
  1. Let $\boldsymbol{X}^* = \boldsymbol{U}^* \boldsymbol{B}^*$.
     Thus $\boldsymbol{x}_k^* = \boldsymbol{U}^* \boldsymbol{b}_k^*$ and so $\boldsymbol{y}_{ik} := |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle| = |\langle \boldsymbol{U}^{*\prime} \boldsymbol{a}_{ik}, \boldsymbol{b}_k^* \rangle|$
  2. If $\boldsymbol{U}^*$ is known, recovering $\boldsymbol{b}_k^*$ is an (easy) r-dimensional standard PR problem
     - ★ needs only $m \geq r$ measurements.
  3. Given an estimate of $\boldsymbol{U}^*$ and of $\boldsymbol{b}_k^*$, we can get an estimate of phase of $\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle$. Updating $\boldsymbol{U}^*$ is then a Least Squares problem
     - ★ can show that for this step mq of order nr^4 suffices.

Key idea of the algorithm: Alt-Min for Phaseless Low Rank Recovery [Nayer, Narayanamurthy, Vaswani, Phaseless PCA, ICML 2019 (this work)]

- Alternating minimization relies on the following key idea:
  1. Let $\boldsymbol{X}^* = \boldsymbol{U}^* \boldsymbol{B}^*$.
     Thus $\boldsymbol{x}_k^* = \boldsymbol{U}^* \boldsymbol{b}_k^*$ and so $\boldsymbol{y}_{ik} := |\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle| = |\langle \boldsymbol{U}^{*\prime} \boldsymbol{a}_{ik}, \boldsymbol{b}_k^* \rangle|$
  2. If $\boldsymbol{U}^*$ is known, recovering $\boldsymbol{b}_k^*$ is an (easy) r-dimensional standard PR problem
     - ★ needs only $m \geq r$ measurements.
  3. Given an estimate of $\boldsymbol{U}^*$ and of $\boldsymbol{b}_k^*$, we can get an estimate of phase of $\langle \boldsymbol{a}_{ik}, \boldsymbol{x}_k^* \rangle$. Updating $\boldsymbol{U}^*$ is then a Least Squares problem
     - ★ can show that for this step mq of order nr^4 suffices.
- Spectral init: compute $\hat{\boldsymbol{U}}^{init}$ as top r eigenvectors of

$$\boldsymbol{Y}_U = \frac{1}{mq} \sum_{k=1}^{q} \sum_{i=1}^{m} \boldsymbol{y}_{ik}^2 \boldsymbol{a}_{ik} \boldsymbol{a}_{ik}{}' \mathbf{1}_{\left\{ \boldsymbol{y}_{ik}^2 \leq \frac{9}{mq} \sum_{ik} \boldsymbol{y}_{ik}^2 \right\}}$$

AltMin-LowRaP: Alt-Min for Phaseless Low Rank Recovery [Nayer, Narayanamurthy, Vaswani, Phaseless PCA, ICML 2019 (this work)]

$\hat{r} \leftarrow$ largest index j for which $\lambda_j(\boldsymbol{Y}_U) - \lambda_n(\boldsymbol{Y}_U) \geq \omega$

$\boldsymbol{U} \leftarrow$ top $\hat{r}$ singular vectors of $\boldsymbol{Y}_U := \frac{1}{mq} \sum_{i,k: \boldsymbol{y}_{ik}^2 \leq \frac{9}{mq} \sum_{ik} \boldsymbol{y}_{ik}^2} \boldsymbol{y}_{ik}^2 \boldsymbol{a}_{ik} \boldsymbol{a}_{ik}{}'$

for $t = 0 : T$ **do**

  $\hat{\boldsymbol{b}}_k \leftarrow \mathrm{RWF}(\{\boldsymbol{y}_k, \boldsymbol{U}' \boldsymbol{a}_{ik}\}, i = 1, 2, \ldots, m)$ for each $k = 1, 2, \cdots, q$

  $\hat{\boldsymbol{X}}^t \leftarrow \boldsymbol{U}\hat{\boldsymbol{B}}$ where $\hat{\boldsymbol{B}} = [\hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \ldots \hat{\boldsymbol{b}}_q]$

  QR decomposition: $\hat{\boldsymbol{B}} \overset{\mathrm{QR}}{=} \boldsymbol{R}_B \boldsymbol{B}$

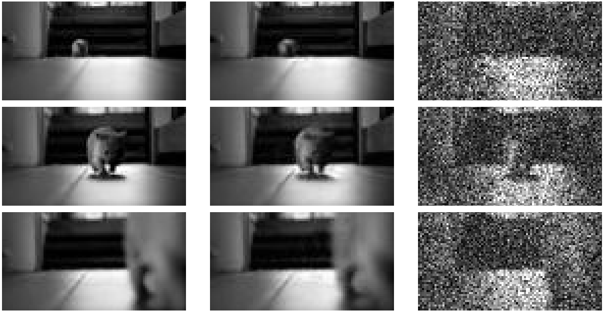
  $\hat{\boldsymbol{c}}_{ik} \leftarrow phase(\langle \boldsymbol{a}_{ik}, \hat{\boldsymbol{x}}_{ik} \rangle), \; i = 1, 2, \ldots, m, \; k = 1, 2, \cdots, q$

  $\hat{\boldsymbol{U}} \leftarrow \arg\min_{\tilde{\boldsymbol{U}}} \sum_{k=1}^{q} \sum_{i=1}^{m} (\hat{\boldsymbol{c}}_{ik} \boldsymbol{y}_{ik} - \boldsymbol{a}_{ik}{}' \tilde{\boldsymbol{U}} \boldsymbol{b}_k)^2$

  QR decomp: $\hat{\boldsymbol{U}} \overset{\mathrm{QR}}{=} \boldsymbol{U}\boldsymbol{R}_U$

end for

RWF: one of two (provably) best standard PR methods



(a) Original (b) AltMinLowRaP (c) RWF

Figure 1: Recovering a real video of a moving mouse (approx low-rank) from simulated $m = 5n$ coded diffraction pattern (CDP) measurements. Showing frames $20, 60, 78$.