Wirtinger Flow Method with Optimal Stepsize for Phase Retrieval

Xue Jiang*, Sreeraman Rajan, and Xingzhao Liu***

* Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (E-mail: {xuejiang, xzhliu}@sjtu.edu.cn).

** Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (E-mail: sreeramanr@sce.carleton.ca).

Abstract

The recently reported Wirtinger flow (WF) algorithm has been demonstrated as a promising method for solving the problem of phase retrieval by applying a gradient descent scheme. An empirical choice of stepsize is suggested in practice. However, this heuristic stepsize selection rule is not optimal. In order to accelerate the convergence rate, we propose an improved WF with optimal stepsize . It is revealed that this optimal stepsize is the solution of a univariate cubic equation with real-valued coefficients. Finding its roots is computationally simple because a closed-form expression exists. Furthermore, compared with obtaining the coefficients of the cubic equation, calculating the gradient is still the leading cost. Therefore, the proposed approach has the same dominant cost as WF in each iteration. Simulation results are provided to validate its efficiency compared to the existing technique.

Introduction

Phase retrieval refers to the task of recovering a signal from phaseless measurements , i.e., reconstruction of a signal given only the magnitudes of its linear measurements. This problem arises in various fields of science and engineering, such as optical imaging, X-ray crystallography, astronomy, and radar, when the phases of the linear transform of the signal are unavailable [1][2]. The phase retrieval problem can be represented as

$$\begin{aligned} &\text{find } \boldsymbol{x} \\ &\text{such that } b_i = |\boldsymbol{a}_i^H \boldsymbol{x}|^2,\ i = 1, \cdots, M \end{aligned} \tag{1}$$

where b_i is the ith phaseless observation of a complex vector $\boldsymbol{x} = [x_1, \cdots, x_N]^T \in \mathbb{C}^N$, and $\boldsymbol{a}_i$ is the ith measurement vector. Note that the problem reduces to solving a system of quadratic equations, which is known as NP hard in general. Furthermore, the equality relationship in (1) cannot hold exactly in the presence of noise.

Wirtinger flow

The following least squares criterion is considered for solving $\boldsymbol{x}$:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \overset{\Delta}{=} \frac{1}{2M} \sum_{i=1}^{M} \left(\left| \boldsymbol{a}_i^H \boldsymbol{x} \right|^2 - b_i \right)^2. \tag{2}$$

- The cost function $f(\boldsymbol{x})$ is not convex. Minimizing non-convex objectives may have very many stationary points.
- Surprisingly, Candès et al. prove that the WF with an initialization using a spectral method converges to the global solution at a geometric rate with high probability provided that the sample size is on the order of $N \log N$ [1].

The WF method is a gradient descent scheme. At the kth iteration, given the current point $\boldsymbol{x}^k$, the estimator can be updated by taking a step along the negative gradient direction. That is,

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - \alpha^k \boldsymbol{g}^k \tag{3}$$

where the gradient $\boldsymbol{g}^k = \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}^*}\big|_{\boldsymbol{x}=\boldsymbol{x}^k}$ can be calculated based on Wirtinger derivatives as $\boldsymbol{g}^k = \boldsymbol{A}^H(\boldsymbol{r} \odot \boldsymbol{A}\boldsymbol{x})$, where $\boldsymbol{A} = \left[\boldsymbol{a}_1^*, \cdots, \boldsymbol{a}_M^*\right]^T \in \mathbb{C}^{M \times N}$ is the measurement matrix, $\boldsymbol{r} = [r_1, \cdots, r_M]^T$ is the error vector with $r_i = \left|\boldsymbol{a}_i^H \boldsymbol{x}\right|^2 - b_i$ and $\alpha^k \in \mathbb{R}_+$ is the current stepsize, which is chosen empirically as

$$\alpha_k = \min\left(1 - e^{-k/k_0}, k_{\max}\right) \tag{4}$$

where k_0 and $k_{\max}$ are two constants with typical values being shown in [1].

Remarks:

- This heuristic stepsize selection rule is not optimal.
- The stepsize controls the convergence rate. The algorithm may converge slowly if the stepsize is too small whereas the algorithm may diverge if the stepsize is too large.
- Our aim is to accelerate the convergence rate of the WF through selecting a more appropriate stepsize.

Wirtinger Flow with Optimal Stepsize (WFOS)

Given $\boldsymbol{x}^k$ and $\boldsymbol{g}^k$ at the kth iteration, the optimal step size can be obtained by solving the line search

$$\boxed{\alpha^k = \arg\min_{\alpha} f(\boldsymbol{x}^k - \alpha \boldsymbol{g}^k).} \tag{5}$$

Note that the cost function $f(\boldsymbol{x}^k - \alpha \boldsymbol{g}^k)$ is a univariate quartic function of α . The optimal stepsize satisfies the following first-order optimality condition:

$$\frac{\mathrm{d} f(\boldsymbol{x}^k - \alpha \boldsymbol{g}^k)}{\mathrm{d}\alpha} = 0$$

which leads to a univariate cubic equation of α given by

$$c_3 \alpha^3 + c_2 \alpha^2 + c_1 \alpha + c_0 = 0 \tag{6}$$

with real-valued constant coefficients $\{c_3, c_2, c_1, c_0\}$. Then the optimal stepsize is the real root associated with the minimum objective value if (6) has three real roots or the minimizer is the unique real root if (6) has a real root and a pair of complex conjugate roots.

Computational Complexity:

- The computational cost for calculating the coefficients $\{c_3, c_2, c_1, c_0\}$ is $\mathcal{O}(M)$.
- The complexity of calculating the roots of a cubic equation is merely $\mathcal{O}(1)$ because a closed-form solution exists.
- Computing the gradient is still the leading cost of the proposed WFOS in each iteration, which is $\mathcal{O}(MN)$.

Simulation Results

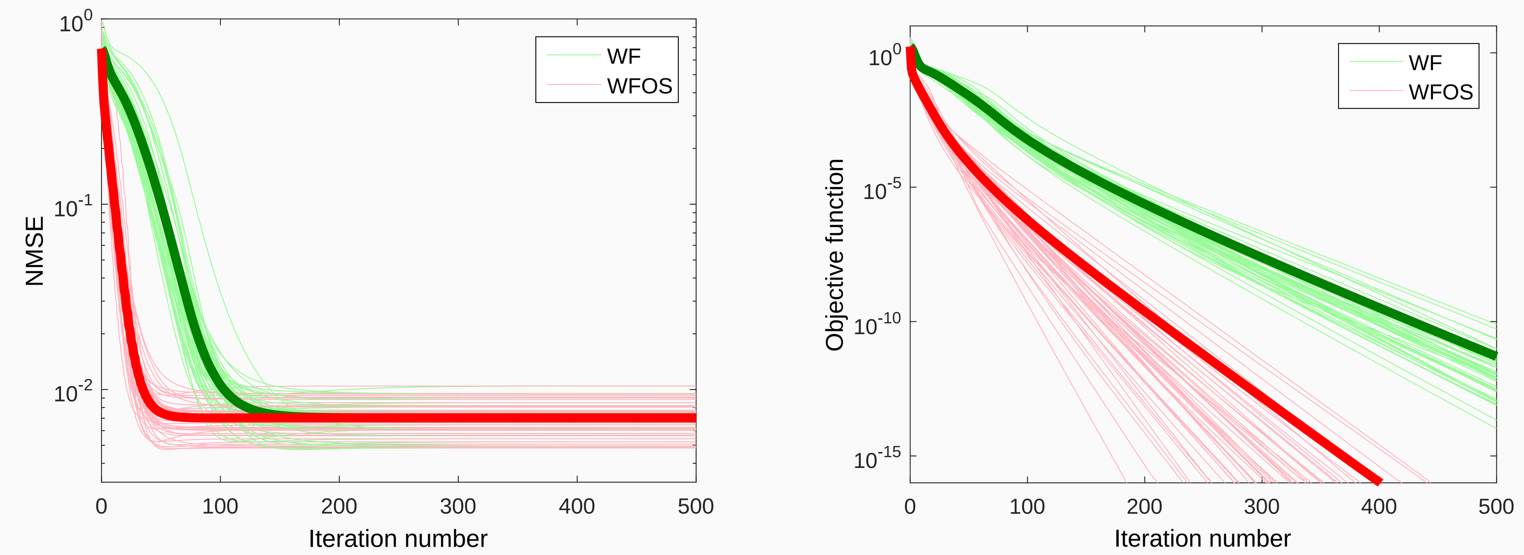

Fig. 1: The convergence behavior of the WFOS compared with WF. The light curves represent the results for each run and thick curves represent the average results over 50 Monte Carlo trials.

Simulation settings:

- Random complex Gaussian signal
- Signal dimension $N = 64$. Observation dimension is six times the signal dimension, i.e, $M = 6N$.
- The same initial value obtained from the spectral method [1].
- The stepsize for WF at the kth iteration is $\alpha_k = \min\left(1 - e^{-k/330}, 0.2\right)$ [1].
- The noise component is sampled from $\mathcal{N}(0, \sigma^2/2) + j\mathcal{N}(0, \sigma^2/2)$ and we have SNR = 20 dB.

Conclusion

- The proposed WFOS for phase retrieval significantly accelerates the convergence rate of WF.
- The optimal stepsize is demonstrated to be the solution of a univariate cubic equation with real-valued coefficients.
- The WFOS has the same leading cost as WF for computation of the gradient in each iteration, which is $\mathcal{O}(MN)$.
- The proposed scheme to obtain the optimal stepsize of WF can also be directly applied to the truncated WF.

References

- E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory,* vol. 61, no.4, pp. 1985–2007, Apr. 2015.
- Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical Imaging: A contemporary overview," *IEEE Signal Process. Mag.,* vol. 32, no. 3, pp. 87–109, May 2015.