Projected Gradient Method for Phase Synchronization

2022年4月22日 9:04

Consider the problem

$$(P) \qquad \hat{z} \in \arg\max \; z^H C z \quad \text{s.t.} \quad z \in \mathbb{T}^n, \qquad \leftarrow C = z^* z^{*H} + \Delta$$

where $\mathbb{T}^n = \{ w \in \mathbb{C}^n : |w_i| = 1 \ \forall i \}$. The PGM for solving (P) is given by

(PGM)

initialize: $z^0 = v_c$

repeat: $w^k \leftarrow \underbrace{z^k}_{\text{current iterate}} + \overbrace{\frac{\alpha}{n}}^{\text{step size}} \underbrace{C z^k}_{\text{gradient}} \quad \longleftarrow \left(I + \frac{\alpha}{n} C \right) z^k$

$$z^{k+1} \leftarrow \frac{w^k}{|w^k|} \quad \longleftarrow \quad \left(\frac{w}{|w|} \right)_j = \begin{cases} \frac{w_j}{|w_j|} & \text{if } w_j \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

Theorem 1 (Estimation Error) — $\underline{\Delta = \sigma W} \overset{\text{whp}}{\Rightarrow} \|\Delta\| = O(\sigma \sqrt{n})$

Suppose that $\|\Delta\| \le \frac{n}{16}$. If $z^0 = v_c$ and $\alpha \ge 2$ in PGM, then

$$d_2(z^{k+1}, z^*) \le \mu^{k+1} d_2(z^0, z^*) + \frac{\nu}{1-\mu} \cdot \underbrace{\frac{8\|\Delta\|}{\sqrt{n}}}_{\ge d_2(v_c, z^*)} \quad \forall k \ge 0,$$

where $\mu = \frac{16(\alpha\|\Delta\| + n)}{(7\alpha + 8)n} < 1$, $\quad \nu = \frac{2\alpha}{7\alpha + 8}$

Remark: Theorem 1 does not guarantee the convergence of PGM

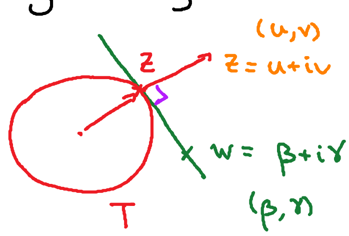
How about the optimization error?

- How to characterize the stationary points of (P)?

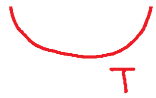
Note that $\mathbb{T}^n$ is a smooth manifold embedded in $\mathbb{C}^n$.

Given $z \in \mathbb{T} \subseteq \mathbb{C}$, the tangent space at z is given by

$$T_z \mathbb{T} = \{ w \in \mathbb{C} : \mathrm{Re}(z \bar{w}) = 0 \}.$$



Then, given $z \in \mathbb{T}^n \subseteq \mathbb{C}^n$, we have



$w = \beta + i\gamma$

(β, γ)

$u\beta + v\gamma = 0$

Then, given $z \in \mathbb{T}^n \subseteq \mathbb{C}^n$, we have

$$\mathcal{T}_z \mathbb{T}^n = \mathcal{T}_{z_1}\mathbb{T} \times \mathcal{T}_{z_2}\mathbb{T} \times \cdots \times \mathcal{T}_{z_n}\mathbb{T}.$$

$$= \{ w \in \mathbb{C}^n : \operatorname{Re}(z_j \bar{w}_j) = 0 \ \forall j \}.$$

It can be easily verified that the projector onto $\mathcal{T}_z\mathbb{T}^n$ is given by

$$\Pi_{\mathcal{T}_z\mathbb{T}^n}(w) = w - \operatorname{Diag}(\operatorname{Re}(z_j \bar{w}_j)) z$$

(Exercise)

Let $f(z) = -z^H C z$ be the objective function to be minimized. The Riemannian gradient of f at z on $\mathbb{T}^n$ is defined as

$$\operatorname{grad} f(z) = \Pi_{\mathcal{T}_z\mathbb{T}^n}(\nabla f(z))$$

$$= 2 \Big(\underbrace{\operatorname{Diag}(\operatorname{Re}((Cz)_j \bar{z}_j)) - C}_{S(z)} \Big) z$$

Then, the first-order optimality condition of (P) is

$$0 = \operatorname{grad}(f(z)) = S(z) z$$

To formulate the second-order optimality condition, we need the Riemannian Hessian, which is the projection of directional derivatives of the Riemannian gradient onto the tangent space:

$$(\operatorname{Hess} f(z))(w) = \Pi_{\mathcal{T}_z\mathbb{T}^n}(D \operatorname{grad} f(z))(w) = \Pi_{\mathcal{T}_z\mathbb{T}^n}(2 S(z) w).$$

Then, the second-order optimality condition of (P) is

$$w^H (\operatorname{Hess} f(z)) w = 2 w^H S(z) w \geq 0 \quad \forall w \in \mathcal{T}_z\mathbb{T}^n$$

Claim: If $z \in \mathbb{T}^n$ satisfies both the first- and second-order

Optimality conditions (e.g., if z is an optimal solution to (P)),

then $$(\mathrm{Diag}(|Cz|) - C)z = 0. \qquad — (*)$$

Consider $\tilde{C} = \frac{n}{\alpha}(I + \frac{\alpha}{n}C) = C + \frac{n}{\alpha}I$. Then, we can also

$C = z^* z^{*H} + \Delta$

write $(*)$ as

$$(\mathrm{Diag}(|\tilde{C}z|) - \tilde{C})z = 0 \qquad — (**)$$

<u>Theorem 2</u>: (Convergence rate of PGM)

Suppose that $\|\Delta\| \le \frac{n^{3/4}}{312}$, $\|\Delta z^*\|_\infty \le \frac{n}{24}$, $\alpha \in [4, \frac{n}{\|\Delta\|})$,

and $z^0 = v_C$. Then, there exist $a > 0$ and $\lambda \in (0,1)$ s.t.

for any optimal solution $\hat{z}$ to (P),

$$g(\hat{z}) - g(z^k) \le (g(\hat{z}) - g(z^0))\lambda^k,$$
$$d_2(z^k, \hat{z}) \le a \cdot (g(\hat{z}) - g(z^0))^{1/2} \lambda^{k/2},$$

where $g(z) = z^H C z$.

The proof consists of two steps. The first is an error bound result:

<u>Proposition 1</u>: Let $\Sigma(z) = \mathrm{Diag}(|\tilde{C}z|) - \tilde{C}$, $\rho(z) = \|\Sigma(z)z\|_2$. Under the assumptions of the theorem, for any $z \in T^n$ satisfying $d_2(z, z^*) \le \sqrt{n}/2$ and for any optimal solution $\hat{z}$ to (P),

$$d_2(z, \hat{z}) \le \frac{8}{n}\rho(z).$$

The second is some properties of PGM:

<u>Proposition 2</u>: Under the assumptions of the theorem, there exist $a_0, a_1, a_2 > 0$ s.t. for any optimal solution $\hat{z}$ to (P), the following hold:

(i) (sufficient ascent) $g(z^{k+1}) - g(z^k) \geq a_0 \cdot \|z^{k+1} - z^k\|_2^2$,

(ii) (cost-to-go estimate) $g(\hat{z}) - g(z^k) \leq a_1 \cdot d_2(z^k, \hat{z})^2$,

(iii) (safeguard) $\rho(z^k) \leq a_2 \|z^{k+1} - z^k\|_2$