On Combining Spectral Initialization and Bayesian GAMP

Junjie Ma

May 31, 2018

Combining spectral initialization and Bayesian GAMP is tricky: we need to properly scale the spectral initial estimate to guarantee the optimality of subsequent iterations. We discuss how it's done in our code.

Recall that the spectral estimator has the following behavor:

$$\boldsymbol{x}^0 \approx \alpha_0 \boldsymbol{x}_\star + \sigma_0 \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}),$$

where α_0 and σ_0 are constants that can be predicted via SE. (Strictly speaking, we can only predict $|\alpha_0|$.) We assume that $\|\boldsymbol{x}_\star\| = \sqrt{N}$. The initial estimate $\boldsymbol{p}^0$ (which can be viewed as an estimate of the variable $\boldsymbol{z}_\star = \boldsymbol{A}\boldsymbol{x}_\star$) is given by

$$\boldsymbol{p}^0 = \boldsymbol{A}\boldsymbol{x}^0 - \text{Onsager}.$$

The existence of the Onsager term ensures that $\boldsymbol{p}^0$ is Gaussian and

$$\begin{bmatrix} \frac{\|\boldsymbol{z}_\star^H\|^2}{M} & \frac{\boldsymbol{z}_\star^H \boldsymbol{p}^0}{M} \\ \frac{(\boldsymbol{p}^0)^H \boldsymbol{z}_\star}{M} & \frac{\|\boldsymbol{p}^0\|^2}{M} \end{bmatrix} \approx \frac{1}{\delta} \begin{bmatrix} \frac{\|\boldsymbol{x}_\star^H\|^2}{N} & \frac{\boldsymbol{x}_\star^H \boldsymbol{x}^0}{N} \\ \frac{(\boldsymbol{x}^0)^H \boldsymbol{x}_\star}{N} & \frac{\|\boldsymbol{x}^0\|^2}{N} \end{bmatrix}$$

To ensure the optimality of the function g_{out} (see GAMP paper), we need

$$\frac{\boldsymbol{z}_\star^H \boldsymbol{p}^0}{M} \approx \frac{\|\boldsymbol{p}^0\|^2}{M}.$$

From the above, this condition transforms into the requirement:

$$\frac{\boldsymbol{x}_\star^H \boldsymbol{x}^0}{N} \approx \frac{\|\boldsymbol{x}^0\|^2}{N}.$$

Hence, for a given $\boldsymbol{x}^0$ that doesn't satisfy the above condition (which is the case for spectral initialization), we need to do the following rescaling:

$$\begin{aligned} \tilde{\boldsymbol{x}}^0 &\leftarrow \boldsymbol{x}^0 \cdot \left(\frac{(\boldsymbol{x}^0)^H \boldsymbol{x}_\star}{\|\boldsymbol{x}^0\|^2} \right) \\ &= \boldsymbol{x}^0 \cdot \frac{\alpha_0^\star}{|\alpha_0|^2 + \sigma_0^2}. \end{aligned}$$

Be careful about the conjugate; the above scaling ensures that $\boldsymbol{x}_\star^H \tilde{\boldsymbol{x}}_0$ is a real positive number. With the above operation, $\tilde{\boldsymbol{x}}_0$ becomes

$$\tilde{\boldsymbol{x}}_0 \approx \frac{|\alpha_0|^2}{|\alpha_0|^2 + \sigma_0^2} \boldsymbol{x}_\star + \frac{|\alpha_0|\sigma_0}{|\alpha_0|^2 + \sigma_0^2} \underbrace{\mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})}_{\boldsymbol{n}}.$$

This can be rewritten as

$$\boldsymbol{x}_\star = \boldsymbol{x}^0 + \underbrace{\frac{-\sigma_0^2}{|\alpha_0|^2 + \sigma_0^2} \boldsymbol{x}^0 + \boldsymbol{n}}_{\boldsymbol{w}}$$

It is not difficult to verify that $\mathbb{E}[X^0 W] = 0$ and

$$\begin{aligned}
\mathbb{E}[|W|^2] &= \mathbb{E}[|X_\star|^2] - \mathbb{E}[|X^0|^2] \\
&= 1 - \left(\frac{|\alpha_0|}{|\alpha_0|^2 + \sigma_0^2} \right)^2 (|\alpha_0|^2 + \sigma_0^2) \\
&= \frac{\sigma_0^2}{|\alpha_0|^2 + \sigma_0^2}.
\end{aligned}$$

Correspondingly, the estimate and distortion terms in terms of $\boldsymbol{z}_\star$ are uncorrected and thus independent (since they are approximately Gaussian). Hence, we can write

$$\boldsymbol{z}_\star = \boldsymbol{p}^0 + \mathcal{CN}\left(0, \frac{1}{\delta} \frac{\sigma_0^2}{|\alpha_0|^2 + \sigma_0^2}\right)$$

Remark 1. *Since we don't know the phase of α_0, there will be a global random phase. Hopefully, this random phase is preserved during the iterations (similar to AMP.A).*