The second-order geometry of short-and-sparse blind deconvolution

Yenson Lau

March 22, 2019

Abstract

These notes describe a set of experiments for interrogating the curvature of nonconvex objective functions for short-and-sparse blind deconvolution (SaS-BD) over the course of optimization.

1 Introduction

In SaS-BD, we observe a length-m signal

$$\boldsymbol{y} \;=\; \iota\boldsymbol{a}_0 \circledast \boldsymbol{x}_0 + \boldsymbol{w},$$

the cyclic convolution between a *short* kernel $\boldsymbol{a}_0 \in \mathbb{S}^{p_0-1}$, zero-padded to length m by $\iota : \mathbb{R}^{m\times p_0}$, and a *sparse* activation map $\boldsymbol{x}_0 \in \mathbb{R}^m$, with possible additive noise $\boldsymbol{w}$. Our goal is to recover $\boldsymbol{a}_0$ and $\boldsymbol{x}_0$ up to some scaling or cyclic shift, as

$$\iota\boldsymbol{a}_0 \circledast \boldsymbol{x}_0 \;\equiv\; \alpha\, s_\ell[\iota\boldsymbol{a}_0] \circledast \alpha^{-1} s_{-\ell}[\boldsymbol{x}_0]$$

for any $\alpha \in \{\mathbb{R}\setminus 0\}$ and $\ell \in \mathbb{Z}$. We refer to these as the scaling and shift symmetries respectively.

A natural approach is to begin by formulating SaS-BD as a nonconvex optimization problem,

$$\min_{\boldsymbol{a}\in\mathbb{S}^{p-1},\boldsymbol{x}} \left[\, \Psi_\lambda(\boldsymbol{a},\boldsymbol{x}) \;\doteq\; \psi(\boldsymbol{a},\boldsymbol{x};\boldsymbol{y}) + \lambda\rho(\boldsymbol{x}) \,\right], \tag{1.1}$$

which minimizes a reconstruction error between $\boldsymbol{y}$ and $\boldsymbol{a}\circledast\boldsymbol{x}$, plus a sparse penalty ρ on $\boldsymbol{x}$. For example, combining a squared error loss with an ℓ_1-norm penalty yields the *bilinear-lasso* (BL) formulation

$$\Psi_\lambda(\boldsymbol{a},\boldsymbol{x}) \;=\; \tfrac{1}{2}\left\|\iota\boldsymbol{a}\circledast\boldsymbol{x}-\boldsymbol{y}\right\|_2^2 + \lambda\left\|\boldsymbol{x}\right\|_1. \tag{BL}$$

A simplified problem, known as the *dropped-quadratic* formulation (DQ), can be obtained when the circulant matrix of $\boldsymbol{a}$, $\boldsymbol{C}_{\boldsymbol{a}}$, satisfies $\boldsymbol{C}_{\boldsymbol{a}}^T\boldsymbol{C}_{\boldsymbol{a}} \simeq \boldsymbol{I}$ over the course of optimization,

$$\Psi_\lambda(\boldsymbol{a},\boldsymbol{x}) \;=\; \tfrac{1}{2}\left\|\boldsymbol{x}\right\|_2^2 - \langle\iota\boldsymbol{a}\circledast\boldsymbol{x},\boldsymbol{y}\rangle + \tfrac{1}{2}\left\|\boldsymbol{y}\right\|_2^2 + \lambda\rho(\boldsymbol{x}), \tag{DQ}$$

$$\rho(\boldsymbol{x}) \;=\; \textstyle\sum_i \left(x_i^2+\delta^2\right)^{\frac{1}{2}}. \tag{1.2}$$

1.1 Regional landscape geometry

When $\boldsymbol{x}_0$ is generic, its effect on the objective landscape can be marginalized via minimization,

$$(1.1) \;\equiv\; \min_{\boldsymbol{a}\in\mathbb{S}^{p-1}} \Big[\, \varphi_\lambda(\boldsymbol{a}) \;\doteq\; \min_{\boldsymbol{x}} \;\Psi_\lambda(\boldsymbol{a},\boldsymbol{x}) \,\Big]. \tag{1.3}$$

As the scale symmetry has already been broken by the sphere, the landscape of φ_λ is primarily driven by shift symmetries after marginalization. Indeed, when $p > p_0$, several cyclic shifts of $\boldsymbol{a}_0$ can potentially be recovered through the variable $\boldsymbol{a}$, each creating a local minimizer on φ_λ. This in turn influences the behavior of φ_λ for regions of the sphere in the vicinity of *multiple shifts*.

Within such regions, the landscape geometry of φ_λ become regularized in a way that appears to be *largely independent* of the specific formulation used for SaS-BD, but finds a relatively simple expression in the (DQ) formulation [KLZW19]. Suppose $\boldsymbol{a} \in \mathcal{S}_\tau \doteq \big\{\boldsymbol{a} : \iota\boldsymbol{a} \in \mathrm{span}\big(\{s_\ell[\boldsymbol{a}_0]\}_{\ell\in\tau}\big)\big\}$, for a few shifts $\tau \subset [m],\ |\tau| \geq 2$, i.e.

$$\iota\boldsymbol{a} \;=\; \textstyle\sum_{\ell\in\tau} \alpha_\ell s_\ell[\iota\boldsymbol{a}_0] \;\in\; \mathbb{S}^{p-1}. \tag{1.4}$$

1.2 Sparsity-coherence tradeoff

2 Experimental outline

References

[KLZW19] Han-Wen Kuo, Yenson Lau, Yuqian Zhang, and John Wright. Geometry and symmetry in short-and-sparse deconvolution. *arXiv preprint arXiv:1901.00256*, 63(7):4497–4520, 2019.