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

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

for any $\alpha \in \{\mathbb{R} \setminus 0\}$ and $l \in \mathbb{Z}$. 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, the *bilinear-lasso* (BL) formulation combines a squared error loss with an ℓ_1-norm penalty,

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

The *dropped-quadratic* formulation (DQ) can be obtained as an approximation to (BL) when the circulant matrix $\boldsymbol{C}_{\boldsymbol{a}}$ of $\boldsymbol{a}$, satisfies $\boldsymbol{C}_{\boldsymbol{a}}^T \boldsymbol{C}_{\boldsymbol{a}} \simeq \boldsymbol{I}$ over the course of optimization,

$$\begin{aligned} \Psi_\lambda(\boldsymbol{a}, \boldsymbol{x}) \;\;&=\;\; \tfrac{1}{2} \left\| \boldsymbol{x} \right\|_2^2 \;-\; \langle \iota \boldsymbol{a} \circledast \boldsymbol{x}, \boldsymbol{y} \rangle \;+\; \tfrac{1}{2} \left\| \boldsymbol{y} \right\|_2^2 + \lambda \rho(\boldsymbol{x}), && \text{(DQ)} \\ \rho(\boldsymbol{x}) \;\;&=\;\; \textstyle\sum_i \left(x_i^2 + \delta^2 \right)^{\frac{1}{2}}. && (1.2) \end{aligned}$$

1.1 Regional landscape geometry

The generic effect of $\boldsymbol{x}_0$ on the objective landscape can be marginalized via minimization,

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

The landscape of φ_λ is then primarily driven by shift symmetries on the sphere. Indeed, each cyclic shift of $\boldsymbol{a}_0$ recoverable through the variable $\boldsymbol{a}$ creates a local minimizer on φ_λ, which in turn influences the behavior of φ_λ for regions of the sphere in the vicinity of *multiple shifts*. This effect appears to be qualitatively *independent of the specific formulation* used for SaS-BD, but is simpler to express in the (DQ) formulation [KLZW19]. Suppose $\boldsymbol{a}$ is near the subspace

$$\mathcal{S}_\tau \doteq \left\{\boldsymbol{a} : \iota\boldsymbol{a} \in \mathrm{span}\left(\{s_l[\iota\boldsymbol{a}_0]\}_{l\in\tau}\right)\right\} \cap \mathbb{S}^{p-1}, \tag{1.4}$$

for a few shifts $\tau \subset \{0,\dots,p-p_0+1\}$. Specifically, this means $\iota\boldsymbol{a} \simeq \sum_{l\in\tau}\alpha_l s_l[\iota\boldsymbol{a}_0]$ is approximately the superposition of a few shifts from $\boldsymbol{a}_0$, the support of each contained in $[p]$. Furthermore, let l_1 and l_2 be the shifts corresponding to first and largest magnitude coefficients in this span. Then the following properties obtain:

Strong convexity near single shifts. If $|\alpha_{l_2}| / |\alpha_{l_1}| \simeq 0$

Negative curvature at balanced points. If $|\alpha_{l_2}| / |\alpha_{l_1}| \simeq 1$

Retraction to subspace.

1.2 Sparsity-coherence tradeoff

1.3 Experimental contributions

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.