Weekly Report 5

Review progress and current progress for writing paper

Yuchen Jin

Dept. of ECE, University of Houston

March 1, 2019

Outline

Outline

Introduction

- **Recent works**:
  - Writing for a new review
    1. note20190227special: Learning the derivation for the convergence of solving inverse problem with deep network. [1, 2]
  - Complete an new article
    1. note20190228special: Derivation for applying the non-negative least square algorithm to solve constrained problem.
  - Progress on writing paper
    1. Have finished the framework of what we are planning for writing. (notes are attached to this slice).
    2. Generate more detail testing estimations.
    3. Propose a possible and slight revision for current work (Need to get more testing results).

Outline

1. Introduction
2. **Review**
3. Paper progress
4. Reference

Introduction

NETT: Solving Inverse Problems with Deep Neural Networks [1]

The regularized form of the problem

$$x_+ \in \arg\min_{x\in\mathbb{D}} \mathcal{T}_{\alpha,\ y_\delta}(x), \qquad \mathcal{T}_{\alpha,\ y_\delta}(x) := \mathcal{D}(\mathbf{F}(x),\ y_\delta) + \alpha\mathcal{R}(\mathbb{V},\ x), \tag{1}$$

The ideal form where the noise is zero

$$x_+ \in \arg\min\{\mathcal{R}(\mathbb{V},\ x) | x \in \mathbb{D} \cap \mathbf{F}(x) = y_0\}. \tag{2}$$

- The paper aims at the analysis of the convergence when
  - the distance function $\mathcal{D}(\cdot)$ is fixed (e.g. the data misfit in our problem).
  - the regularization term is realized by dictionary learning or deep network.
    - Dictionary learning: $\mathcal{R}(\mathbb{V},\ \mathbf{x}) = |\mathbf{V\Phi x}|_q^q = \sum_{\lambda=1}^N v_\lambda |\boldsymbol{\varphi}_\lambda^T \mathbf{x}|^q$.
    - Deep network (CNN): $\mathcal{R}(\mathbb{V},\ x) = \psi(\mathbf{\Phi}(\mathbb{V},\ x))$.

The formulation of the network

$$\Phi(\mathbb{V},\ x) := (\sigma_L \circ \mathbb{V}_L \circ \sigma_{L-1} \circ \mathbb{V}_{L-1} \circ \cdots \circ \sigma_1 \circ \mathbb{V}_1)(x), \quad (3)$$

$$\mathbb{V}_l := \mathbb{A}_l(x) + b_l, \quad (4)$$

$$\mathbb{A}_{l,\ \lambda}(x) := \sum_{\mu=1}^{N^{(l-1)}} \mathcal{K}^{(l)}_{\mu,\ \lambda}(x_\mu). \quad (5)$$

- $\sigma_l \circ \mathbb{V}_l$ represents "a layer of the network", where
  - $\sigma_l(\cdot)$ is the non-linear activating function.
  - $\mathbb{V}_l(\cdot)$ is viewed as an affine linear operator.
  - $\mathbb{A}_l(\cdot)$ is defined by convolutions that maps the previous layer into the current one. b_l is the bias to different channels.

Introduction

NETT: Solving Inverse Problems with Deep Neural Networks [1]

Conditions for the weak convergence

- The regularizer $\mathcal{R}(\mathbb{V},\ \mathbf{x})$ satisfies that:
  - For layer l, we have $\mathbb{A}_l(x)$ is bounded linear, i.e. $\exists\ c_l,\ \|x\| \leqslant c_l\|\mathbb{A}(x)\|$;
  - θ_l is a weekly continuous and coercive function;
  - ψ is a lower semi-continuous and coercive function.
- The data consistency term should satisfied that
  - For a $\tau \geqslant 1$, we have for any y_1, y_2, y, $\mathcal{D}(y_1,\ y_2) \leqslant \tau(\mathcal{D}(y_1,\ y) + \mathcal{D}(y,\ y_2))$;
  - For any y, y_0, $\mathcal{D}(y,\ y_0) = 0 \iff y = y_0$;
  - For any $(y_k)_{k\in\mathbb{N}}$, $y_k \to y \implies \mathcal{D}(y_k,\ y) \to 0$;
  - $\mathcal{D}(\mathbf{F}(x),\ y)$ is sequentially lower semi-continuous with respect to x.

- Hence we could find from the conditions of $\mathcal{R}$:
  - $\mathcal{R}$ is weakly and sequentially lower semi-continuous;
  - Since $\mathcal{R}$ is coercive, for all $t > 0, \alpha > 0$ and y, we have that $\{x|\mathcal{T}_{\alpha,\ y}(x) \leqslant t\}$ is sequentially weakly (closed and bounded).

Introduction

NETT: Solving Inverse Problems with Deep Neural Networks [1]

Theorem for weak convergence

- For all $\alpha > 0$ and y, $\exists\ \mathcal{T}_{\alpha,\ y}(x)$.
- If $y_k \to y$ and $x_k \in \arg\min \mathcal{T}_{\alpha,\ y}(x)$, there exists the weak accumulation points $(x_k)_{k\in\mathbb{N}}$ which is derived from the minimizer $\mathcal{T}_{\alpha,\ y}(x)$.
- Consider $y = \mathbf{F}(x)$, we have $(y_k)_{k\in\mathbb{N}}$ that $\mathcal{D}(y_k,\ y) + \mathcal{D}(y,\ y_k) \leqslant \delta_k$, where $(\delta_k)_{k\in\mathbb{N}} \to 0$. Suppose that x_k is k^{th} iterative solution, i.e. $x_k \in \arg\min \mathcal{T}_{\alpha,\ \delta_k}(x,\ y_k)$. Then we could choose α that

$$\lim_{\delta\to 0} \alpha(\delta) = \lim_{\delta\to 0} \frac{\delta}{\alpha(\delta)} = 0. \tag{6}$$

  - $(x_k)_{k\in\mathbb{N}}$ has at least one weak accumulation point xx_+.
  - Any weakly convergent subsequence $(x_k)_{k\in\mathbb{N}}$ satisfies $\mathcal{R}(\mathbb{V},\ x_{k(n)}) \to \mathcal{R}(\mathbb{V},\ x_+)$.
  - If a the solution of (2) is unique, then we know $(x_k) \to x_+$.

Introduction

NETT: Solving Inverse Problems with Deep Neural Networks [1]

Absolute Bregman distance

$$\mathcal{B}_{\mathcal{F}}(\tilde{x},\ x) = |\mathcal{F}(\tilde{x}) - \mathcal{F}(x) - \mathcal{F}'(x)(\tilde{x} - x)|. \tag{7}$$

- Absolute Bregman distance shows that when the local part around x of the function could be fitted by linear function, $\mathcal{B}_{\mathcal{F}}(x + \Delta x,\ x) = 0$.
- If for any x, $\mathcal{B}_{\mathcal{F}}(x + \Delta x,\ x) > 0$, the function would be totally non-linear.
- With the total non-linearity, the weak convergence would be used to derive the strong convergence, where
  - The solution of weak convergence is $x_{k(n)}$, which satisfies $\mathcal{R}(\mathbb{V},\ x_{k(n)}) \to \mathcal{R}(\mathbb{V},\ x_+)$.
  - $\mathcal{B}_{\mathcal{F}}(x_{k(n)},\ x_+) \to 0$.
  - Hence we get $\|x_{k(n)} - x_+\| \to 0$.

Outline

Paper progress

- In the following part I would use ☑ to show the completed parts.
  - ☑ Introduction
  - ☑ Related works
  - Problem background
    - Industrial background
    - Data
    - Geosteering inverse problem
    - Challenges
  - Proposed method
    - Bayesian view of our problem
    - Proposed network
      - ☐ Train
      - ☐ Test
  - Testing results
- It seems that our core idea has been produced by [3].

Outline

Reference I

- H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, "NETT: Solving Inverse Problems with Deep Neural Networks," *arXiv e-prints*, p. arXiv:1803.00092, Feb 2018.
- E. Ekici, S. Jafari, M. Caldas, and T. Noiri, "Weakly λ-continuous functions," *Novi Sad J. Math*, vol. 38, no. 2, pp. 47–56, 2008.
- J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, no. 12, p. 124007, nov 2017. [Online]. Available: https://doi.org/10.1088%2F1361-6420%2Faa9581

Thank you for Listening

It's time for Q & A