Estimation in Rotationally Invariant Generalized Linear Models via Approximate Message Passing

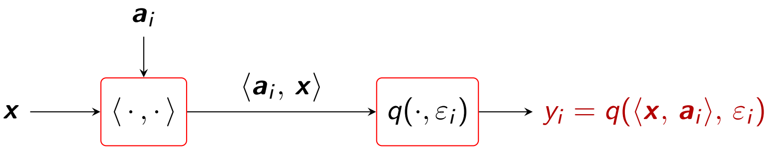
Ramji Venkataramanan
University of Cambridge

Kevin Kögler
ISTA

Marco Mondelli
ISTA

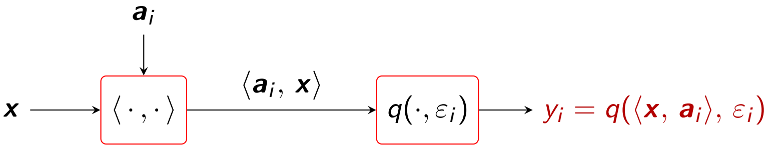
ICML 2022

Generalized Linear Models



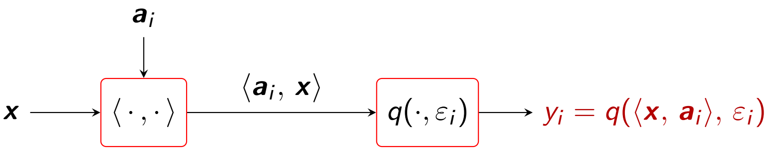
- GOAL: Estimate signal $\boldsymbol{x} \in \mathbb{R}^d$ from observations $y_1, \ldots, y_n$
- Known sensing vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ and output function q

Examples



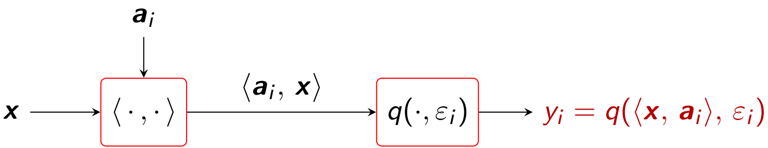
- Linear model $y_i = \langle \boldsymbol{x}, \boldsymbol{a}_i \rangle + \varepsilon_i$

Examples



- Linear model $y_i = \langle \boldsymbol{x}, \boldsymbol{a}_i \rangle + \varepsilon_i$
- Phase retrieval $y_i = |\langle \boldsymbol{x}, \boldsymbol{a}_i \rangle|^2 + \varepsilon_i$

Examples



- Linear model $y_i = \langle \boldsymbol{x}, \boldsymbol{a}_i \rangle + \varepsilon_i$
- Phase retrieval $y_i = |\langle \boldsymbol{x}, \boldsymbol{a}_i \rangle|^2 + \varepsilon_i$
- 1-bit compressed sensing $y_i = \text{sign}(\langle \boldsymbol{x}, \boldsymbol{a}_i \rangle)$

Rotationally Invariant Model

$$\boldsymbol{A} = \begin{bmatrix} \longleftarrow & \boldsymbol{a}_1^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \boldsymbol{a}_n^T & \longrightarrow \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad \boldsymbol{y} = q(\boldsymbol{A}\boldsymbol{x}, \varepsilon)$$

Rotationally invariant $\boldsymbol{A}$

- SVD of $\boldsymbol{A} = \boldsymbol{O}\boldsymbol{\Lambda}\boldsymbol{Q}^T$
- $\boldsymbol{O}, \boldsymbol{Q}$ uniformly random orthogonal matrices
- Arbitrary singular values $\boldsymbol{\Lambda}$
- More general than Gaussian $\boldsymbol{A}$, can capture complex correlation structure in the data

Rotationally Invariant Model

$$\boldsymbol{A} = \begin{bmatrix} \longleftarrow & \boldsymbol{a}_1^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \boldsymbol{a}_n^T & \longrightarrow \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad \boldsymbol{y} = q(\boldsymbol{A}\boldsymbol{x}, \varepsilon)$$

Rotationally invariant $\boldsymbol{A}$

- SVD of $\boldsymbol{A} = \boldsymbol{O}\boldsymbol{\Lambda}\boldsymbol{Q}^T$
- $\boldsymbol{O}, \boldsymbol{Q}$ uniformly random orthogonal matrices
- Arbitrary singular values $\boldsymbol{\Lambda}$
- More general than Gaussian $\boldsymbol{A}$, can capture complex correlation structure in the data

High-dimensional setting: $d, n \to \infty$ and $\frac{n}{d} \to \delta \in \mathbb{R}$

Main contributions

1. RI-GAMP: Approximate Message Passing algorithm for rotationally invariant models

- Iteratively produces estimates $\boldsymbol{x}^t$, for $t \geq 1$
- Can be tailored to take advantage of prior info about signal

Main contributions

1. RI-GAMP: Approximate Message Passing algorithm for rotationally invariant models

   - Iteratively produces estimates $\boldsymbol{x}^t$, for $t \geq 1$
   - Can be tailored to take advantage of prior info about signal

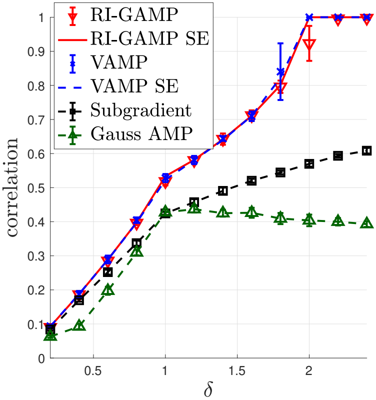
2. Rigorous characterization of RI-GAMP iterates in high-dimensional setting via **state evolution**

Gives precise formulas for

Limiting MSE : $\lim_{d\to\infty} \frac{1}{d}\|\boldsymbol{x}-\boldsymbol{x}^t\|^2$, Correlation : $\lim_{d\to\infty} \frac{|\langle \boldsymbol{x}, \boldsymbol{x}^t\rangle|}{\|\boldsymbol{x}\|\|\boldsymbol{x}^t\|}$

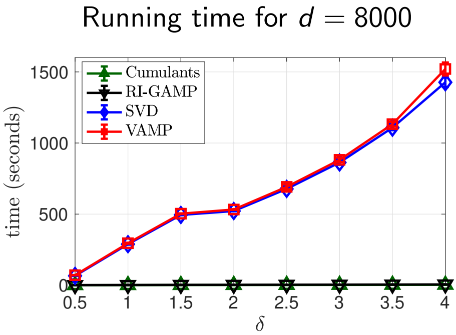
Example: 1-bit compressed sensing

$$\boldsymbol{y} = \text{sign}(\boldsymbol{A}\boldsymbol{x})$$



$\boldsymbol{x}$ with Rademacher prior, singular values of $\boldsymbol{A} \sim \sqrt{6}\,\mathrm{Beta}(1,2)$

Comparison with Vector AMP



VAMP conjectured to be Bayes-optimal in some settings, but requires expensive SVD computation: $\mathcal{O}(d^3)$

RI-GAMP matches VAMP performance, but much faster: $\mathcal{O}(d^2)$

Idea of proof

We design an **auxiliary AMP** that

- uses knowledge of true signal $\boldsymbol{x}$
- is an instance of abstract AMP recursion for rotationally invariant matrices [Zhong et al. '21], can be analyzed

Show that state evolution for RI-GAMP matches auxiliary AMP and that the iterates are close

Take-away

$$\boldsymbol{A} = \begin{bmatrix} \longleftarrow & \boldsymbol{a}_1^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \boldsymbol{a}_n^T & \longrightarrow \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad \boldsymbol{y} = q(\boldsymbol{A}\boldsymbol{x}, \varepsilon)$$

- Novel AMP estimator for rotationally invariant GLMs
- Sharp asymptotic performance guarantees via State Evolution
- Performance matching VAMP, but with lower complexity
  VAMP conjectured to be Bayes-optimal in some settings