ICASSP 2018, Calgary, Canada

Linear Quantization by Effective Resistance Sampling

Yining Wang
Carnegie Mellon University

Joint work with Aarti Singh

QUANTIZED LINEAR SENSING

- The linear model: $y = X\beta_0$
  - ***X***: n by p "design" matrix, with full knowledge
  - ***y***: n-dim vector, the sensing result
  - $\boldsymbol{\beta_0}$: p-dim unknown signal to be recovered

QUANTIZED LINEAR SENSING

- The linear model: $y = X\beta_0$
- The quantized sensing problem:
  - Measurements of y cannot be made in arbitrary precision
  - A total of k bits allocated to each measurement y_i
  - Each y_i rounded to the nearest integer with k_i binary bits.

$$\widetilde{y}_i = 2^{-(k_i-1)} \cdot \text{round}\left[2^{k_i-1}\frac{y_i}{M}\right]$$

QUANTIZED LINEAR SENSING

- The linear model: $y = X\beta_0$
- The quantized sensing problem:
  - Measurements of y cannot be made in arbitrary precision
  - A total of k bits allocated to each measurement y_i
  - Each y_i rounded to the nearest integer with k_i binary bits.

$$\widetilde{y}_i = 2^{-(k_i-1)} \cdot \text{round}\left[2^{k_i-1}\frac{y_i}{M}\right]$$

Range of y

Quantized linear sensing

- The linear model: $y = X\beta_0$
- The quantized sensing problem:
  - Measurements of y cannot be made in arbitrary precision
- Example applications:
  - Brain activity measurements: total signal strength limited
  - Distributed sensing: signal communication limited

QUANTIZED LINEAR SENSING

- The linear model: $y = X\beta_0$
- The quantized sensing problem:

$$\widetilde{y}_i = 2^{-(k_i - 1)} \cdot \mathrm{round}\left[2^{k_i - 1} \frac{y_i}{M}\right]$$

- Question: **how to allocate measurement bits to achieve the best statistical efficiency?**

Dithering

- "Dithering": $\widetilde{y}_i = 2^{-(k_i-1)} \cdot \text{round}\left[2^{k_i-1}\left(\frac{y_i}{M} + \delta_i\right)\right]$
  - Introducing artificial noise for independent statistical error
  - Equivalent model: $\widetilde{y}_i = \langle x_i, \beta_0 \rangle + \varepsilon_i$

$$\mathbb{E}[\varepsilon_i] = 0 \qquad \mathbb{E}[\varepsilon_i^2] \leq 4^{-(k_i+1)} M^2$$

Dithering

uniform noise between two values

- "Dithering": $\widetilde{y}_i = 2^{-(k_i-1)} \cdot \mathrm{round}\left[2^{k_i-1}\left(\frac{y_i}{M} + \delta_i\right)\right]$
  - Introducing artificial noise for independent statistical error
  - Equivalent model: $\widetilde{y}_i = \langle x_i, \beta_0 \rangle + \varepsilon_i$

$$\mathbb{E}[\varepsilon_i] = 0 \qquad \mathbb{E}[\varepsilon_i^2] \leq 4^{-(k_i+1)} M^2$$

WEIGHTED OLS

- "Dithering": $\widetilde{y}_i = 2^{-(k_i-1)} \cdot \mathrm{round}\left[2^{k_i-1}\left(\frac{y_i}{M} + \delta_i\right)\right]$
  $= \langle x_i, \beta_0 \rangle + \varepsilon_i \qquad \mathbb{E}[\varepsilon_i^2] \leq 4^{-(k_i+1)} M^2$
- Weighted Ordinary Least Squares (OLS)
  $\widehat{\beta}_{\boldsymbol{k}} = (X^\top W X)^{-1} X^\top W \widetilde{y}$

WEIGHTED OLS

- "Dithering": $\widetilde{y}_i = 2^{-(k_i-1)} \cdot \text{round}\left[2^{k_i-1}\left(\frac{y_i}{M} + \delta_i\right)\right]$

  $= \langle x_i, \beta_0 \rangle + \varepsilon_i \qquad \mathbb{E}[\varepsilon_i^2] \leq 4^{-(k_i+1)} M^2$

- Weighted Ordinary Least Squares (OLS)

  $\widehat{\beta}_{\boldsymbol{k}} = (X^\top W X)^{-1} X^\top W \widetilde{y}$

  $W = \text{diag}(w_1, w_2, \cdots, w_n)$

  $= \text{diag}(4^{k_1+1}, 4^{k_2+1}, \cdots, 4^{k_n+1})$

Weighted OLS

- Weighted Ordinary Least Squares (OLS)

$$\widehat{\beta}_{\boldsymbol{k}} = (X^\top W X)^{-1} X^\top W \widetilde{y}$$

$$\mathbb{E}\|\widehat{\beta}_{\boldsymbol{k}} - \beta_0\|_2^2 \leq M^2 \cdot \mathrm{tr}\left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top\right]^{-1}$$

- Optimal quantization:

$$\min_{\boldsymbol{k}} \mathrm{tr}[X^\top W X]^{-1} \qquad s.t.\;\; k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

Weighted OLS

- Weighted Ordinary Least Squares (OLS)

$$\widehat{\beta}_{\boldsymbol{k}} = (X^\top W X)^{-1} X^\top W \widetilde{y}$$

$$X^\top W X$$

$$\mathbb{E}\|\widehat{\beta}_{\boldsymbol{k}} - \beta_0\|_2^2 \le M^2 \cdot \mathrm{tr}\left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top\right]^{-1}$$

- Optimal quantization:

$$\min_{\boldsymbol{k}} \mathrm{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \le k, \quad k_i \in \mathbb{N}$$

Weighted OLS

- Weighted Ordinary Least Squares (OLS)

$$\widehat{\beta}_{\boldsymbol{k}} = (X^\top W X)^{-1} X^\top W \widetilde{y} \qquad X^\top W X$$

$$\mathbb{E}\|\widehat{\beta}_{\boldsymbol{k}} - \beta_0\|_2^2 \leq M^2 \cdot \mathrm{tr}\left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top\right]^{-1}$$

- Optimal quantization:

Combinatorial… hard!

$$\min_{\boldsymbol{k}} \mathrm{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

Continuous relaxation

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t.\ \ k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

- Still a challenging problem…
  - Non-convexity of objectives!

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \quad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- Still a challenging problem…
  - Non-convexity of objectives!

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \quad s.t. \quad k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- A re-formulation:

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top \right]^{-1} \quad s.t. \quad \sum_{i=1}^{n} k_i \leq k$$

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} \quad s.t. \quad \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \quad k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

$k_i \in \mathbb{R}^+$

- A re-formulation:

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top \right]^{-1} \qquad s.t. \quad \sum_{i=1}^{n} k_i \leq k$$

Convex objective

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} \qquad s.t. \quad \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$k_i \in \mathbb{R}^+$

- A re-formulation:

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top \right]^{-1} \qquad s.t. \;\; \sum_{i=1}^{n} k_i \leq k$$

Convex objective

Non-convex feasible set

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} \qquad s.t. \;\; \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- A re-formulation:

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} \qquad s.t. \;\; \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

$$\downarrow$$

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} + \lambda \left[\sum_{i=1}^{n} \log_4(w_i) - (n-k) \right]$$

Continuous relaxation

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \quad k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- A re-formulation:

$$\min \operatorname{tr}\left[\sum_{i=1}^{n} w_i x_i x_i^\top\right]^{-1} \qquad s.t. \quad \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

Convex objective

$$\min \operatorname{tr}\left[\sum_{i=1}^{n} w_i x_i x_i^\top\right]^{-1} + \lambda\left[\sum_{i=1}^{n} \log_4(w_i) - (n-k)\right]$$

CONTINUOUS RELAXATION

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \quad k_1 + \cdots + k_n \leq k, \quad k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- A re-formulation:

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} \qquad s.t. \quad \sum_{i=1}^{n} \log_4(w_i) - 1 \leq k$$

Convex objective

concave objective

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} w_i x_i x_i^\top \right]^{-1} + \lambda \left[\sum_{i=1}^{n} \log_4(w_i) - (n-k) \right]$$

Continuous relaxation

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- A re-formulation:
  - DC (Difference of Convex functions) programming:

$$\min \operatorname{tr}\left[\sum_{i=1}^{n} w_i x_i x_i^\top\right]^{-1} - \lambda \left[-\sum_{i=1}^{n} \log_4(w_i) + (n-k)\right]$$

Rounding / Sparsification

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$k_i \in \mathbb{R}^+$

- How to obtain integral solutions? “Sparsify” $\boldsymbol{k}$
  - Idea 1: round to the nearest integer
  - Problem: might cause objective to increase significantly

Rounding / Sparsification

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \mathrm{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$$k_i \in \mathbb{R}^+$$

- How to obtain integral solutions? "Sparsify" $\boldsymbol{k}$
  - Idea 2: **simple** sampling
    - *Sample* i *from the distribution normalized by* **k**
    - $k(i) = k(i) + 1$
  - Problem: slow convergence (require large budget *k*)

Rounding / Sparsification

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t. \;\; k_1 + \cdots + k_n \leq k, \;\; k_i \in \mathbb{N}$$

$k_i \in \mathbb{R}^+$

- How to obtain integral solutions? "Sparsify" $\boldsymbol{k}$
  - Idea 3: **effective resistance** sampling

$$t \sim p_t \propto 4^{k_t+1} \ell_t$$

  - Advantage: fast convergence (*k* independent of condition numbers of ***X*** or ***W***.

Rounding / Sparsification

- Continuously relaxed optimal quantization:

$$\min_{\boldsymbol{k}} \operatorname{tr}[X^\top W X]^{-1} \qquad s.t.\;\; k_1 + \cdots + k_n \le k, \;\; k_i \in \mathbb{N}$$

$k_i \in \mathbb{R}^+$

- How to obtain integral solutions? "Sparsify" $\boldsymbol{k}$
  - Idea 3: **effective resistance** sampling

$$t \sim p_t \propto 4^{k_t+1}\ell_t$$

Effective resistance:

$$\ell_t = x_t^\top [W^*]^{-1} x_t$$

  - Advantage: fast convergence (k independent of condition numbers of $\boldsymbol{X}$ or $\boldsymbol{W}$.

Open questions

- Most important question: how to solve (continuous)

$$\min \operatorname{tr} \left[\sum_{i=1}^{n} 4^{k_i+1} x_i x_i^\top \right]^{-1} \qquad s.t. \;\; \sum_{i=1}^{n} k_i \leq k$$

- Some ideas:
  - Is the objective quasi-convex or directional convex?
  - Are local minima also global, or approximately global?
    - *Escaping saddle point methods?*
  - Are there adequate **convex** relaxations?

Thank you!
Questions