Generalization Error Analysis of Quantized Compressive Learning

Xiaoyun Li  Ping Li
Department of Statistics, Rutgers University
Cognitive Computing Lab, Baidu Research USA
Random Projection (RP) Method

- Data matrix $X \in \mathbb{R}^{n \times d}$, normalized to unit norm (all samples on unit sphere).
- Save storage by $k$ random projections: $X_R = X \times R$, with $R \in \mathbb{R}^{d \times k}$, a random matrix with i.i.d. $N(0, 1)$ entries $\implies X_R \in \mathbb{R}^{n \times k}$.
- J-L lemma: approximate distance preservation $\implies$ Many applications: clustering, classification, compressed sensing, dimensionality reduction, etc..
- “Projection+quantization”: apply (entry-wise) scalar quantization function $Q(\cdot)$ by $X_Q = Q(X_R)$. More storage saving.
- More applications: MaxCut, SimHash, 1-bit compressive sensing, etc.
Compressive Learning + Quantization

- We can apply learning models to \( \left( \frac{1}{\sqrt{k}} X_R, Y \right) \mapsto \) learning in the projected space \( S_R \)!

- This is called **compressive learning**. It can be shown that learning in the projected space is able to provide satisfactory performance, while substantially reduce the computational cost, especially for high-dimensional data.

- We go one step further: learning with \( \left( \frac{1}{\sqrt{k}} Q(X_R), Y \right) \mapsto \) learning in the quantized projected space \( S_Q \)!

- This is called **quantized compressive learning**. Very practical in applications with data compression.
We provide generalization error bounds (of a test sample) on quantized compressive learning models:
- Nearest neighbor classifier
- Linear classifier (logistic regression, linear SVM, etc.)
- Linear regression

**Applications**: we identify the factors that affect the generalization performance of each model, which gives recommendations on the choice of quantizer $Q$ in practice.

Some experiments are conducted to verify the theory.
Some Background on Quantization

A $b$-bit quantizer $Q_b$ separates the real line into $M = 2^b$ regions. Suppose signal $X \sim f$.

- **Distortion**: $D_{Q_b} = E[(Q_b(X) - X)^2] \iff$ minimized by Lloyd-Max (LM) quantizer.
- **Maximal gap** of $Q$ on interval $[a, b]$: the largest gap between two consecutive boarders of $Q$ on $[a, b]$.
- **Inner product estimator**: $\hat{\rho}_{Q_b}(x_1, x_2) = \frac{Q_b(x_1^TR)Q_b(R^Tx_2)}{k}$.

**Definition: Debiased Variance**

Denote the space of expectation of estimator $\hat{\rho}_{Q_b}$ as $\mathcal{E}$. If there exists a map $g : [-1, 1] \to \mathcal{E}$, the **debiased estimator** is defined by applying the inverse map $\hat{\rho}_{Q_b}^{db} = g^{-1}(\hat{\rho})$ to correct for the bias. The variance of $\hat{\rho}_{Q_b}^{db}$ is called the **debiased variance**.
We are interested in the risk of a classifier $h$, $\mathcal{L}(h) = E[\mathbb{1}\{h(x) \neq y\}]$.

Assume $(x, y) \sim \mathcal{D}$, with conditional probability $\eta(x) = P(y = 1|x)$. Bayes classifier $h^*(x) = \mathbb{1}\{\eta(x) > 1/2\}$ has the minimal risk.

$h_Q(x) = y_Q^{(1)}$, where $(x_Q^{(1)}, y_Q^{(1)})$ is the sample and label of nearest neighbor of $x$ in the quantized space $S_Q$.

**Theorem: Generalization of 1-NN Classifier**

Suppose $(x, y)$ is a test sample. $Q$ is a uniform quantizer with $\triangle$ between boarders and maximal gap $g_Q$. Under some technical conditions and with some constants $c_1, c_2$, with high probability,

$$E_{X,Y}[\mathcal{L}(h_Q(x))] \leq 2\mathcal{L}(h^*(x)) + c_1\left(\frac{\triangle}{g_Q}\sqrt{\frac{1 + \omega}{1 - \omega}}\right)^{k+1}(ne)^{-\frac{1}{k+1}}\sqrt{k} + \frac{c_2\triangle\sqrt{k}}{\sqrt{1 - \omega}}.$$
Quantized Compressive 1-NN Classifier: Asymptotics

Theorem: Asymptotic Error of 1-NN Classifier

Let the cosine estimator $\hat{\rho}_Q = \frac{Q(x_1^T R)Q(R^T x_2)}{k}$, assume $\forall x_1, x_2$, $E[\hat{\rho}_Q(x_1, x_2)] = \alpha \rho_{x_1, x_2}$ for some $\alpha > 0$. As $k \to \infty$, we have

$$E_{X,Y,R}[\mathcal{L}(h_Q(x))] \leq E_{X,Y}[\mathcal{L}(h_S(x))] + r_k,$$

$$r_k = E\left[\sum_{i:x_i \in \mathcal{G}} \Phi\left(\sqrt{k}(\cos(x, x_i) - \cos(x, x^{(1)})) \sqrt{\xi_{x,x_i}^2 + \xi_{x,x^{(1)}}^2 - 2 \text{Corr}(\hat{\rho}_Q(x, x_i), \hat{\rho}_Q(x, x^{(1)})) \xi_{x,x_i} \xi_{x,x^{(1)}}}\right)\right],$$

with $\xi_{x,y}^2 / k$ the debiased variance of $\hat{\rho}_Q(x, y)$ and $\mathcal{G} = X / x^{(1)}$. $\mathcal{L}(h_S(x))$ is the risk of data space NN classifier, and $\Phi(\cdot)$ is the CDF of $N(0, 1)$.

- Under mild conditions, smaller debiased variance around $\cos(x, x^{(1)})$ leads to smaller generalization error.
Quantized Compressive Linear Classifier

- $H$ separates the space by a hyper-plane: $H(x) = 1\{h^T x > 0\}$.
- ERM classifiers: $\hat{H}(x) = 1\{\hat{h}^T x > 0\}$, $\hat{H}_Q(x) = 1\{\hat{h}_Q^T Q(R^T x) > 0\}$.

**Theorem: Generalization of linear classifier**

Under some technical conditions, with probability $(1 - 2\delta)$,

$$
Pr[\hat{H}_Q(x) \neq y] \leq \hat{L}_{(0,1)}(S, \hat{h}) + \frac{1}{\delta n} \sum_{i=1}^{n} f_{k,Q}(\rho_i) + C_{k,n,\delta},
$$

where $f_{k,Q}(\rho_i) = \Phi(-\frac{\sqrt{k} |\rho_i|}{\xi_{\rho_i}})$, with $\rho_i$ the cosine between training sample $x_i$ and ERM classifier $\hat{h}$ in the data space, and $\xi_{\rho_i}^2/k$ the debiased variance of $\hat{\rho}_Q = \frac{Q(x_1^T R)Q(R^T x_2)}{k}$ at $\rho_i$.

- Small debiased variance around $\rho = 0$ lowers the bound.
- Solution: Lloyd-Max (LM) quantizer achieves minimal debiased variance at 0, and thus should be the first choice.
Quantized Compressive Least Squares (QCLS)

- Fixed design: \( Y = X^T \beta + \epsilon \), with \( x_i \) fixed, \( \epsilon \) i.i.d. \( N(0, \gamma) \)
- \( L(\beta) = \frac{1}{n} E_Y [\| Y - X \beta \|^2] \), \( L_Q(\beta_Q) = \frac{1}{n} E_{Y,R} [\| Y - Q(XR)\beta_Q \|^2] \).
- \( \hat{L}(\beta) = \frac{1}{n} \| Y - X \beta \|^2 \), \( \hat{L}_Q(\beta_Q) = \frac{1}{n} \| Y - \frac{1}{\sqrt{k}} Q(XR)\beta_Q \|^2 \). (given \( R \))

**Theorem: Generalization of QCLS**

Let \( \hat{\beta}^* = \operatorname{argmin}_{\beta \in \mathbb{R}^d} \hat{L}(\beta) \) and \( \hat{\beta}_Q^* = \operatorname{argmin}_{\beta \in \mathbb{R}^k} \hat{L}_Q(\beta) \). Let \( \Sigma = X^T X / k, \ k < n \).

\( D_Q \) is the distortion of \( Q \). Then we have

\[
E_{Y,R} [L_Q(\hat{\beta}_Q^*)] - L(\beta^*) \leq \gamma \frac{k}{n} + \frac{1}{k} \| \beta^* \|^2_{\Omega}, \tag{1}
\]

where \( \Omega = \left[ \frac{\xi_{2,2} - 1 + D_Q}{(1 - D_Q)^2} - 1 \right] \Sigma + \frac{1}{1 - D_Q} I_d \), with \( \| w \|_{\Omega} = \sqrt{w^T \Omega w} \) the Mahalanobis norm.

- Smaller distortion decreased the error bound.
- Solution: the distortion optimal LM quantizer is the first choice.
Conclusions From Theory

- **1-NN classification**: we should choose the quantizer with small debiased variance of inner product estimate \( \hat{\rho}_Q = \frac{Q(R^T x)^T Q(R^T x^{(1)})}{k} \) at around \( \rho = \cos(x, x^{(1)}) \). Here, \( x^{(1)} \) is the nearest neighbor of \( x \). \( \Rightarrow \) Normalizing the quantized random projections may help, see ref Xiaoyun Li and Ping Li, Random Projections with Asymmetric Quantization, NeurIPS 2019.

- **Linear classification**: we should choose the quantizer with small debiased variance of inner product estimate \( \hat{\rho}_Q = \frac{Q(R^T x)^T Q(R^T y)}{k} \) at around \( \rho = 0 \). \( \Rightarrow \) First choice: Lloyd-Max quantizer.

- **Linear regression**: we should choose the quantizer with small distortion \( D_Q \). \( \Rightarrow \) First choice: Lloyd-Max quantizer.
### Some Experiments

<table>
<thead>
<tr>
<th>Dataset</th>
<th># samples</th>
<th># features</th>
<th># classes</th>
<th>Mean 1-NN ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASEHOCK</td>
<td>1993</td>
<td>4862</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>orlraws10P</td>
<td>100</td>
<td>10304</td>
<td>10</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Mean 1-NN ρ** is the estimated $\cos(x, x^{(1)})$ from training set.

---

**Figure 1**: Empirical debiased variance of three quantizers.
Quantized Compressive 1-NN Classification

Claim: smaller debiased variance at $\rho = \cos(x, x^{(1)})$ is better.

Figure 2: Quantized compressive 1-NN classification.

- Target $\rho$ should be around:
  - BASEHOCK: 0.6, where red quantizer has largest debiased variance
  - Orlraws10P: 0.9, where red quantizer has smallest debiased variance
Quantized Compressive Linear SVM

Claim: smaller debiased variance at $\rho = 0$ is better.

![Graph showing test accuracy vs number of projections for BASEHOCK and orlraws10P datasets.](image)

**Figure 3:** Quantized compressive linear SVM.

- At $\rho = 0$, red quantizer has much larger debiased variance than others $\Rightarrow$ Lowest test accuracy on both datasets.
QCLS: Simulation

Claim: smaller distortion is better.

Figure 4: Test MSE of QCLS.

- The order of test error exactly agrees with the order of distortion!