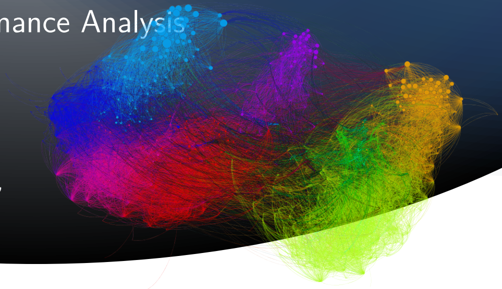
Graphs in Machine Learning

Analysis of Online SSL

Quantization Error and Performance Analysis

Michal Valko

Inria & ENS Paris-Saclay, MVA



Partially based on material by: Branislav Kveton,
il Belkin, Jerry Zhu

Online SSL with Graphs: Analysis

Want to bound $\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t]-y_t)^2$

What can we guarantee?

Online SSL with Graphs: Analysis

Want to bound $\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t]-y_t)^2$

What can we guarantee?

Three sources of error

Online SSL with Graphs: Analysis

Want to bound $\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t]-y_t)^2$

What can we guarantee?

Three sources of error

- generalization error — if all data: $(\hat{f}_{s,t} - y_t)^2$
- online error — data only incrementally: $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$
- quantization error — memory limitation: $(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2$

Online SSL with Graphs: Analysis

Want to bound $\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t]-y_t)^2$

What can we guarantee?

Three sources of error

- generalization error — if all data: $(\hat{f}_{s,t}-y_t)^2$
- online error — data only incrementally: $(\hat{f}_{so,t}[t]-\hat{f}_{s,t})^2$
- quantization error — memory limitation: $(\hat{f}_{soq,t}[t]-\hat{f}_{so,t}[t])^2$

All together:

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t]-y_t)^2 \leq \frac{9}{2N}\sum_{t=1}^{N}(\hat{f}_{s,t}-y_t)^2 + \frac{9}{2N}\sum_{t=1}^{N}(\hat{f}_{so,t}[t]-\hat{f}_{s,t})^2 + \frac{9}{2N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[$$

Online SSL with Graphs: Analysis

Bounding transduction error $(\hat{f}_{s,t} - y_t)^2$

Online SSL with Graphs: Analysis

Bounding transduction error $(\hat{f}_{s,t} - y_t)^2$

If all labeled examples l are i.i.d., $c_l = 1$ and $c_l \gg c_u$, then

$$R(\hat{\ell}^\star) \;\leq\; \widehat{R}(\hat{\ell}^\star) + \underbrace{\beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta + 4)}_{\text{transductive error } \Delta_T(\beta, n_l, \delta)}$$

$$\beta \;\leq\; 2\left[\frac{\sqrt{2}}{\gamma_g + 1} + \sqrt{2n_l}\frac{1 - c_u}{c_u}\frac{\lambda_M(\mathbf{L}) + \gamma_g}{\gamma_g^2 + 1}\right]$$

holds with the probability of $1 - \delta$, where

$$R(\hat{\ell}^\star) = \frac{1}{N}\sum_t (\hat{f}_{s,t} - y_t)^2 \quad \text{and} \quad \widehat{R}(\hat{\ell}^\star) = \frac{1}{n_l}\sum_{t \in l}(\hat{f}_{s,t} - y_t)^2$$

Online SSL with Graphs: Analysis

Bounding **transduction error** $(\hat{f}_{s,t} - y_t)^2$

If all labeled examples l are i.i.d., $c_l = 1$ and $c_l \gg c_u$, then

$$R(\hat{\ell}^\star) \;\leq\; \widehat{R}(\hat{\ell}^\star) + \underbrace{\beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta + 4)}_{\text{transductive error } \Delta_T(\beta, n_l, \delta)}$$

$$\beta \;\leq\; 2\left[\frac{\sqrt{2}}{\gamma_g + 1} + \sqrt{2n_l}\frac{1 - c_u}{c_u}\frac{\lambda_M(\mathbf{L}) + \gamma_g}{\gamma_g^2 + 1}\right]$$

holds with the probability of $1 - \delta$, where

$$R(\hat{\ell}^\star) = \frac{1}{N}\sum_t (\hat{f}_{s,t} - y_t)^2 \quad \text{and} \quad \widehat{R}(\hat{\ell}^\star) = \frac{1}{n_l}\sum_{t \in l}(\hat{f}_{s,t} - y_t)^2$$

How should we set γ_g?

Online SSL with Graphs: Analysis

Bounding transduction error $(\hat{f}_{s,t} - y_t)^2$

If all labeled examples l are i.i.d., $c_l = 1$ and $c_l \gg c_u$, then

$$R(\hat{\ell}^\star) \leq \widehat{R}(\hat{\ell}^\star) + \underbrace{\beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta + 4)}_{\text{transductive error } \Delta_T(\beta, n_l, \delta)}$$

$$\beta \leq 2\left[\frac{\sqrt{2}}{\gamma_g + 1} + \sqrt{2n_l}\frac{1 - c_u}{c_u}\frac{\lambda_M(\mathbf{L}) + \gamma_g}{\gamma_g^2 + 1}\right]$$

holds with the probability of $1 - \delta$, where

$$R(\hat{\ell}^\star) = \frac{1}{N}\sum_t (\hat{f}_{s,t} - y_t)^2 \quad \text{and} \quad \widehat{R}(\hat{\ell}^\star) = \frac{1}{n_l}\sum_{t \in l}(\hat{f}_{s,t} - y_t)^2$$

How should we set γ_g? Want $\Delta_T(\beta, n_l, \delta) = o(1)$

Online SSL with Graphs: Analysis

Bounding transduction error $(\hat{f}_{s,t} - y_t)^2$

If all labeled examples l are i.i.d., $c_l = 1$ and $c_l \gg c_u$, then

$$R(\hat{\ell}^\star) \leq \widehat{R}(\hat{\ell}^\star) + \underbrace{\beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta + 4)}_{\text{transductive error } \Delta_T(\beta, n_l, \delta)}$$

$$\beta \leq 2\left[\frac{\sqrt{2}}{\gamma_g + 1} + \sqrt{2n_l}\frac{1 - c_u}{c_u}\frac{\lambda_M(\mathbf{L}) + \gamma_g}{\gamma_g^2 + 1}\right]$$

holds with the probability of $1 - \delta$, where

$$R(\hat{\ell}^\star) = \frac{1}{N}\sum_t (\hat{f}_{s,t} - y_t)^2 \quad \text{and} \quad \widehat{R}(\hat{\ell}^\star) = \frac{1}{n_l}\sum_{t \in l}(\hat{f}_{s,t} - y_t)^2$$

How should we set γ_g? Want $\Delta_T(\beta, n_l, \delta) = o(1) \to \beta = o\left(n_l^{-1/2}\right)$

Online SSL with Graphs: Analysis

Bounding **transduction error** $(\hat{f}_{s,t} - y_t)^2$

If all labeled examples l are i.i.d., $c_l = 1$ and $c_l \gg c_u$, then

$$R(\hat{\ell}^\star) \;\leq\; \widehat{R}(\hat{\ell}^\star) + \underbrace{\beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta + 4)}_{\text{transductive error } \Delta_T(\beta, n_l, \delta)}$$

$$\beta \;\leq\; 2\left[\frac{\sqrt{2}}{\gamma_g + 1} + \sqrt{2n_l}\frac{1 - c_u}{c_u}\frac{\lambda_M(\mathbf{L}) + \gamma_g}{\gamma_g^2 + 1}\right]$$

holds with the probability of $1 - \delta$, where

$$R(\hat{\ell}^\star) = \frac{1}{N}\sum_t (\hat{f}_{s,t} - y_t)^2 \quad \text{and} \quad \widehat{R}(\hat{\ell}^\star) = \frac{1}{n_l}\sum_{t \in l}(\hat{f}_{s,t} - y_t)^2$$

How should we set γ_g? Want $\Delta_T(\beta, n_l, \delta) = o(1) \rightarrow \beta = o\left(n_l^{-1/2}\right)$

$\rightarrow \gamma_g = \Omega\left(n_l^{1+\alpha}\right)$ for any $\alpha > 0$

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n\times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})}$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1}$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n\times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n\times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$$

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \le \|A\mathbf{v}\|_2 \le \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \le \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \le \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \le \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2$$

Online SSL with Graphs: Analysis

Bounding **online error** $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n\times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \le \|A\mathbf{v}\|_2 \le \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \le \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \le \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \le \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2 \le \|\hat{f}_{so}[t] - \hat{f}_s\|_2^2$$

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_2^2 \leq \left(\frac{2\sqrt{n_l}}{\gamma_g + 1}\right)^2$$

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n\times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_2^2 \leq \left(\frac{2\sqrt{n_l}}{\gamma_g + 1}\right)^2$$

Again, how should we set γ_g?

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_2^2 \leq \left(\frac{2\sqrt{n_l}}{\gamma_g + 1}\right)^2$$

Again, how should we set γ_g? If we want $O\left(n_l^{-1/2}\right)$?

Online SSL with Graphs: Analysis

Bounding online error $(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2$

Idea: If $\mathbf{L}$ and $\hat{L}_o$ are regularized, then HFSs get closer together.

since they get closer to zero

Recall $\hat{\ell} = (\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})^{-1}\mathbf{y}$, where $\mathbf{Q} = \mathbf{L} + \gamma_g \mathbf{I}$

and also $\mathbf{v} \in \mathbb{R}^{n \times 1}$, $\lambda_m(A)\|\mathbf{v}\|_2 \leq \|A\mathbf{v}\|_2 \leq \lambda_M(A)\|\mathbf{v}\|_2$

$$\|\hat{\ell}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\lambda_m(\mathbf{C}^{-1}\mathbf{Q} + \mathbf{I})} = \frac{\|\mathbf{y}\|_2}{\frac{\lambda_m(\mathbf{Q})}{\lambda_M(\mathbf{C})} + 1} \leq \frac{\sqrt{n_l}}{\gamma_g + 1}$$

Difference between offline and online solutions:

$$(\hat{f}_{so,t}[t] - \hat{f}_{s,t})^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_\infty^2 \leq \|\hat{f}_{so}[t] - \hat{f}_s\|_2^2 \leq \left(\frac{2\sqrt{n_l}}{\gamma_g + 1}\right)^2$$

Again, how should we set γ_g? If we want $O\left(n_l^{-1/2}\right)$? Then $\gamma_g = \Omega\left(n_l^{3/4}\right)$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t]-\hat{f}_{so,t}[t]\right)^2$

How are the quantized and full solution different?

$$\hat{\ell}^\star = \min_{\hat{\ell}\in\mathbb{R}^N} (\hat{\ell}-\mathbf{y})^\mathsf{T}\mathbf{C}(\hat{\ell}-\mathbf{y}) + \hat{\ell}^\mathsf{T}\mathbf{Q}\hat{\ell}$$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

How are the quantized and full solution different?

$$\hat{\ell}^\star = \min_{\hat{\ell} \in \mathbb{R}^N} (\hat{\ell} - \mathbf{y})^\mathsf{T} \mathbf{C} (\hat{\ell} - \mathbf{y}) + \hat{\ell}^\mathsf{T} \mathbf{Q} \hat{\ell}$$

In $\mathbf{Q}$!

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

How are the quantized and full solution different?

$$\hat{\ell}^{\star} = \min_{\hat{\ell} \in \mathbb{R}^N} (\hat{\ell} - \mathbf{y})^{\mathsf{T}} \mathbf{C} (\hat{\ell} - \mathbf{y}) + \hat{\ell}^{\mathsf{T}} \mathbf{Q} \hat{\ell}$$

In $\mathbf{Q}$! $\hat{K}_o$ (online) vs. $\tilde{K}$ (quantized)

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

How are the quantized and full solution different?

$$\hat{\ell}^\star = \min_{\hat{\ell} \in \mathbb{R}^N} (\hat{\ell} - \mathbf{y})^\mathsf{T} \mathbf{C} (\hat{\ell} - \mathbf{y}) + \hat{\ell}^\mathsf{T} \mathbf{Q} \hat{\ell}$$

In $\mathbf{Q}$! $\hat{K}_o$ (online) vs. $\tilde{K}$ (quantized)

We have: $\hat{\ell}_o = (\mathbf{C}^{-1}\hat{K}_o + \mathbf{I})^{-1}\mathbf{y}$ vs. $\hat{\ell}_{oq} = (\mathbf{C}^{-1}\tilde{K} + \mathbf{I})^{-1}\mathbf{y}$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

How are the quantized and full solution different?

$$\hat{\ell}^\star = \min_{\hat{\ell}\in\mathbb{R}^N} (\hat{\ell} - \mathbf{y})^\mathsf{T}\mathbf{C}(\hat{\ell} - \mathbf{y}) + \hat{\ell}^\mathsf{T}\mathbf{Q}\hat{\ell}$$

In $\mathbf{Q}$! $\hat{K}_o$ (online) vs. $\tilde{K}$ (quantized)

We have: $\hat{\ell}_o = (\mathbf{C}^{-1}\hat{K}_o + \mathbf{I})^{-1}\mathbf{y}$ vs. $\hat{\ell}_{oq} = (\mathbf{C}^{-1}\tilde{K} + \mathbf{I})^{-1}\mathbf{y}$

With linear algebra we get

$$\|\hat{\ell}_{oq} - \hat{\ell}_o\|_2 \leq \frac{\sqrt{n_l}}{c_u\gamma_g^2}\|\hat{K}_{oq} - \hat{K}_o\|_F$$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

The quantization error depends on $\|\hat{K}_{oq} - \hat{K}_o\|_F = \|\hat{L}_{oq} - \hat{L}_o\|_F$.

When can we keep $\|\hat{L}_{oq} - \hat{L}_o\|_F$ under control?

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

The quantization error depends on $\|\hat{K}_{oq} - \hat{K}_o\|_F = \|\hat{L}_{oq} - \hat{L}_o\|_F$.

When can we keep $\|\hat{L}_{oq} - \hat{L}_o\|_F$ under control?

Charikar guarantees **distortion** error of at most $RR/(R-1)$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

The quantization error depends on $\|\hat{K}_{oq} - \hat{K}_o\|_F = \|\hat{L}_{oq} - \hat{L}_o\|_F$.

When can we keep $\|\hat{L}_{oq} - \hat{L}_o\|_F$ under control?

Charikar guarantees **distortion** error of at most $RR/(R-1)$

For what kind of data $\{\mathbf{x}_i\}_{i=1,\dots,n}$ is the distortion small?

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

The quantization error depends on $\|\hat{K}_{oq} - \hat{K}_o\|_F = \|\hat{L}_{oq} - \hat{L}_o\|_F$.

When can we keep $\|\hat{L}_{oq} - \hat{L}_o\|_F$ under control?

Charikar guarantees **distortion** error of at most $RR/(R-1)$

For what kind of data $\{\mathbf{x}_i\}_{i=1,\ldots,n}$ is the distortion small?

Assume manifold $\mathcal{M}$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

The quantization error depends on $\|\hat{K}_{oq} - \hat{K}_o\|_F = \|\hat{L}_{oq} - \hat{L}_o\|_F$.

When can we keep $\|\hat{L}_{oq} - \hat{L}_o\|_F$ under control?

Charikar guarantees **distortion** error of at most $RR/(R-1)$

For what kind of data $\{\mathbf{x}_i\}_{i=1,\dots,n}$ is the distortion small?

Assume manifold $\mathcal{M}$

- all $\{\mathbf{x}_i\}_{i\geq 1}$ lie on a smooth d-dimensional compact $\mathcal{M}$
- with boundary of bounded geometry Def. 11 of Hein **hein2007graph**
  - has finite volume V
  - has finite surface area A
  - should not intersect itself
  - should not fold back onto itself

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

Bounding $\|\hat{L}_{oq} - \hat{L}_o\|_F$ when $\mathbf{x}_i \in \mathcal{M}$

Consider k-sphere packing* of radius r with centers contained in $\mathcal{M}$. *only the centers are packed, not necessarily the entire ball

If k is large $\rightarrow r <$ **injectivity radius** of $\mathcal{M}$ **hein2007graph**

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

Bounding $\|\hat{L}_{oq} - \hat{L}_o\|_F$ when $\mathbf{x}_i \in \mathcal{M}$

Consider k-sphere packing* of radius r with centers contained in $\mathcal{M}$. *only the centers are packed, not necessarily the entire ball

If k is large $\rightarrow r <$ **injectivity radius** of $\mathcal{M}$ **hein2007graph** and $r < 1$:

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

Bounding $\|\hat{L}_{oq} - \hat{L}_o\|_F$ when $\mathbf{x}_i \in \mathcal{M}$

Consider k-sphere packing* of radius r with centers contained in $\mathcal{M}$. *only the centers are packed, not necessarily the entire ball

If k is large $\rightarrow r <$ **injectivity radius** of $\mathcal{M}$ **hein2007graph** and $r < 1$:

$$r < \left(\frac{V + Ac_{\mathcal{M}}}{kc_d}\right)^{1/d} = \mathcal{O}\left(\frac{1}{k^{1/d}}\right)$$

r-packing is a $2r$-covering:

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t]-\hat{f}_{so,t}[t]\right)^2$

Bounding $\|\hat{L}_{oq} - \hat{L}_o\|_F$ when $\mathbf{x}_i \in \mathcal{M}$

Consider k-sphere packing* of radius r with centers contained in $\mathcal{M}$. *only the centers are packed, not necessarily the entire ball

If k is large $\rightarrow r <$ **injectivity radius** of $\mathcal{M}$ **hein2007graph** and $r < 1$:

$$r < \left(\frac{V + Ac_{\mathcal{M}}}{kc_d}\right)^{1/d} = \mathcal{O}\left(\frac{1}{k^{1/d}}\right)$$

r-packing is a $2r$-covering:

$$\max_{i=1,\ldots,N} \|\mathbf{x}_i - \mathbf{c}\|_2 \le R\frac{R}{R-1} \le 2\mathcal{O}\left(k^{-1/d}\right) = \mathcal{O}\left(k^{-1/d}\right)$$

But what about $\|\hat{L}_{oq} - \hat{L}_o\|_F$?

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

If similarity is M-Lipschitz

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

If similarity is M-Lipschitz, $\mathbf{L}$ is normalized

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

If similarity is M-Lipschitz, $\mathbf{L}$ is normalized, $\forall i, j \ \frac{\sqrt{\hat{D}_{o,ii}\hat{D}_{o,jj}}}{N} > c_{min}$

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

If similarity is M-Lipschitz, $\mathbf{L}$ is normalized, $\forall i, j \; \frac{\sqrt{\hat{D}_{o,ii}\hat{D}_{o,jj}}}{N} > c_{min}$

$$\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(MR^2/c_{min}) = \mathcal{O}(k^{-2/d}).$$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

If similarity is M-Lipschitz, $\mathbf{L}$ is normalized, $\forall i, j \ \frac{\sqrt{\hat{D}_{o,ii}\hat{D}_{o,jj}}}{N} > c_{min}$

$$\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(MR^2/c_{min}) = \mathcal{O}(k^{-2/d}).$$

Are the assumptions reasonable?

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d})$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t]-\hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2 \leq \frac{n_l}{c_u^2\gamma_g^4}\|\hat{L}_{oq} - \hat{L}_o\|_F^2$$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2 \leq \frac{n_l}{c_u^2\gamma_g^4}\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \frac{n_l}{c_u^2\gamma_g^4}$$

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2 \leq \frac{n_l}{c_u^2\gamma_g^4}\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \frac{n_l}{c_u^2\gamma_g^4}$$

With properly setting γ_g, e.g., $\gamma_g = \Omega(n_l^{3/8})$, we can have

Online SSL with Graphs: Analysis

Bounding quantization error $\left(\hat{f}_{soq,t}[t]-\hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2 \leq \frac{n_l}{c_u^2\gamma_g^4}\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \frac{n_l}{c_u^2\gamma_g^4}$$

With properly setting γ_g, e.g., $\gamma_g = \Omega(n_l^{3/8})$, we can have

$$\frac{1}{N}\sum_{t=1}^{N}\left(\hat{f}_{soq,t}[t]-y_t\right)^2 = \mathcal{O}\left(n_l^{-1/2}\right).$$

Online SSL with Graphs: Analysis

Bounding **quantization error** $\left(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t]\right)^2$

We showed $\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \mathcal{O}(k^{-2/d}) = \mathcal{O}(1)$.

$$\frac{1}{N}\sum_{t=1}^{N}(\hat{f}_{soq,t}[t] - \hat{f}_{so,t}[t])^2 \leq \frac{n_l}{c_u^2\gamma_g^4}\|\hat{L}_{oq} - \hat{L}_o\|_F^2 \leq \frac{n_l}{c_u^2\gamma_g^4}$$

With properly setting γ_g, e.g., $\gamma_g = \Omega(n_l^{3/8})$, we can have

$$\frac{1}{N}\sum_{t=1}^{N}\left(\hat{f}_{soq,t}[t] - y_t\right)^2 = \mathcal{O}\left(n_l^{-1/2}\right).$$

What does that mean?

Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA

https://misovalko.github.io/mva-ml-graphs.html