Circulant Binary Embedding (F. Yu et al, ICML'14)

Circulant Binary Embedding

It refers to methods for embedding vectors in $\mathcal{S}^{d-1}$ into vertices in the Hamming cube of dimension d with

$$h(\boldsymbol{x}) = \mathsf{sgn}\left(\boldsymbol{G}_c^\top \boldsymbol{D}\boldsymbol{x}\right),$$

where $\boldsymbol{D} \in \mathbb{R}^{d\times d}$ is a diagonal matrix with a Rademacher sequence and $\boldsymbol{G}_c \in \mathbb{R}^{d\times d}$ is a circulant matrix defined as

$$\boldsymbol{G}_c = \begin{pmatrix} g_1 & g_d & \cdots & g_3 & g_2 \\ g_2 & g_1 & \cdots & g_4 & g_3 \\ \vdots & g_2 & g_1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots & g_d \\ g_d & g_{d-1} & \cdots & g_2 & g_1 \end{pmatrix}.$$

Bit Complexities of Circulant Binary Embedding

Table: Comparison of the analysis for BE with unstructured projection and circulant projection, where ϵ is a distortion rate and n is #points

Methods	Bit Complexity	Conditions
Unstructured BE	$\mathcal{O}\left(\epsilon^{-2}\log n\right)$	-
Our analysis	$\mathcal{O}\left(\epsilon^{-2}\log n\right)$	*small infinity norm*
Existing works (Near-optimal)	$\mathcal{O}\left(\epsilon^{-3}\log n\right)$ or $\mathcal{O}\left(\epsilon^{-2}\log^2 n\right)$	*small infinity norm*

Contribution

- We develop a non-trivial extension of existing analysis to achieve the optimal bit complexity of CBE.
- Our analysis is well matched to the original implementation of CBE and empirical justification.