Randomized Algorithms 2021A – Lecture 8 (second part)
Oblivious Subspace Embedding*

Robert Krauthgamer

1 Oblivious Subspace Embedding

Embedding an entire subspace: In some situations (like regression, as we will see soon), we want a guarantee for a whole subspace, which has infinitely many points.

Observe that a linear subspace $V \subset \mathbb{R}^n$ of dimension d can be described as the column space of $A \in \mathbb{R}^{n\times d}$, i.e., $V = \{Ax : x \in \mathbb{R}^d\}$.

A good way to think about the next definition is that we will solve a problem in $\mathbb{R}^n$ involving an unknown d-dimensional subspace, by reducing the problem to dimension $s = s(n, d, \varepsilon, \delta)$. Thus, we want s (the number of rows in S) to be as small as possible.

Definition: A random matrix $S \in \mathbb{R}^{s\times n}$ is called an (ε, δ, d)-*Oblivious Subspace Embedding (OSE)* if

$$\forall A \in \mathbb{R}^{n\times d}, \qquad \Pr_S \Big[\forall x \in \mathbb{R}^d,\ \|SAx\| \in (1 \pm \varepsilon)\|Ax\|\Big] \geq 1 - \delta.$$

We next show that it is easy to construct OSE using JLT.

Exer: Show that the OSE property is preserved under right-muliplication by a matrix with orthonormal columns, as follows. If $S \in \mathbb{R}^{s\times n}$ is an (ϵ, δ, d)-OSE matrix, and $U \in \mathbb{R}^{n\times r}$ is a matrix with orthonormal columns, then SU is an $(\epsilon, \delta, \min(r, d))$-OSE matrix (for the space $\mathbb{R}^r$).

Theorem: Let $S \in \mathbb{R}^{s\times n}$ be an (ε, δ, b)-JLT for $\varepsilon < 1/4$. Then S is also an $(O(\varepsilon), \delta, \frac{\ln b}{\ln(1/\varepsilon)})$-OSE.

Remark: To produce OSE for dimension d, we should set in this theorem $d = \frac{\ln b}{\ln(1/\varepsilon)}$, i.e., $b = (1/\varepsilon)^d$, which we can achieve using a Gaussian matrix with $s = O(\varepsilon^{-2}\log(b/\delta)) = O(\varepsilon^{-2}(d\log\frac{1}{\varepsilon} + \log\frac{1}{\delta}))$ rows. A direct construction with sparse columns (and thus fast matrix-vector multiplication) was shown by [Cohen, 2016].

Proof: Was seen in class. The main idea is to use the JLT guarantee on a (3ε)-net N of the unit sphere in $\mathbb{R}^d$, then represent arbitrary $x \in \mathbb{R}^d$ as an infinite (but converging) sum $x = \sum_{i=0}^{\infty} x_i$,

*These notes summarize the material covered in class, usually skipping proofs, details, examples and so forth, and possibly adding some remarks, or pointers. The exercises are for self-practice and need not be handed in. In the interest of brevity, most references and credits were omitted.

where each x_i is a (scalar) multiple of a net point, and finally use the triangle inequality. We used the next exercise, whose proof is based on volume arguments.

Exer: Show that one can construct a γ-net N of size $|N| \leq (1+2/\gamma)^d \leq (3/\gamma)^d$.

Hint: Let B_r be a ball of radius $r > 0$ in $\mathbb{R}^d$. Then the volume of B_{2r} is bigger than that of B_r by a factor of 2^d.

Remark: It is possible to get a better bound by employing a 1/2-net (instead of ε-net) and expanding $\|SAx\|^2$ including cross terms.