HW 3

Due on **Wednesday, April 16, 2025 at 9am CST**

598GAA: Geometric Approximation Algorithms, Spring 2025 Version: **1.1**

HW Guidelines: As in previous homework. Solutions must be self contained, and containing all relevant details.

In the following $f(n) = \widetilde{O}(g(n))$ implies that $f(n) = O\Big(g(n) \cdot (\log n + 1/\varepsilon)^{O(1)}\Big)$ where ε is the relevant approximation parameter.

7 (100 PTS.) Some JL questions.

7.A. BOXES CAN BE SEPARATED. (Easy) Let A and B be two axis parallel boxes that are interior disjoint. Prove that there is always an axis parallel hyperplane that separates the interior of the two boxes.

7.B. BRUNN-MINKOWSKI INEQUALITY, SLIGHT EXTENSION.
Prove the following claim.

Corollary 3.1. *For A and B compact sets in $\mathbb{R}^n$, we have for any $\lambda \in [0,1]$ that*

$$\mathrm{Vol}(\lambda A + (1-\lambda)B) \geq \mathrm{Vol}(A)^{\lambda}\,\mathrm{Vol}(B)^{1-\lambda}.$$

7.C. PROJECTIONS ARE CONTRACTIONS
(Easy) Let F be a k-dimensional affine subspace, and let $P_F : \mathbb{R}^{\mathsf{d}} \to F$ be the projection that maps every point $x \in \mathbb{R}^{\mathsf{d}}$ to its nearest neighbor on F. Prove that P_F is a contraction (i.e., 1-Lipschitz). Namely, for any $p, q \in \mathbb{R}^{\mathsf{d}}$, we have that $\|P_F(p) - P_F(q)\| \leq \|p - q\|$.

8 (100 PTS.) Some reweighting questions.

8.A. Prove Theorem 7.2.6 from the reweighting notes.

8.B. SPANNING TREE WITH RELATIVE CROSSING NUMBER. Let P be a set of n points in the plane. For a line ℓ, let $w^+(\ell)$ (resp., $w^-(\ell)$) be the number of points of $pSet$ lying above (resp., below or on) ℓ, and define the *weight* of ℓ, denoted by $w\ell$, to be $\min(w^+(\ell), w^-(\ell))$.
Show that one can construct a spanning tree T for P such that any line ℓ crosses $O\big(\sqrt{\omega\ell}\log(n/\omega\ell)\big)$ edges of T.

9 (100 PTS.) Not too many near-neighbor queries.

Let P be a set of n points in a metric space $M = (Z, d)$ (i.e., $P \subseteq Z$), and let $\varepsilon > 0$ be some parameter. Assume that for any two points $p, q \in X$, one can compute their distance $d(p,q)$ in the metric space M in $O(1)$ time. For a point $p \in X$, $\mathsf{n}(p)$ be the distance between p and its nearest neighbor in $P \setminus \{p\}$ (computing this quantity naively takes $O(n)$ time).

Assume you are given a data-structure that answer approximate near-neighbor queries on any prespecified subset $X \subseteq P$ (***APLEB query***). Specifically, given $\varepsilon \in (0,1)$, $r > 0$, and X, assume that one can construct in $O(|X|/\varepsilon)$ time/space, a data-structure $\mathcal{DS}(X, r, \varepsilon)$, such that given a query point q in $O(1)$ time, the data-structure returns one of the following results:

(I) **Short**: This happens only if $d(q, X) = \min_{p \in X} d(q, p) \leq (1 + eps)r$. In this case, the data-structure returns a point $p \in X$ such that $d(p, q) \leq (1 + eps)r$.

(II) **Long**: This happens only if $d(q, X) > r$.

(III) $d(q, X) \in [r, (1 + \varepsilon)r]$ the data-structure is allowed to return either answer.

9.A. (10 PTS.) Prove that for any $\psi \in (0, 1)$, if $d(q, P) > 4\text{diameter}(P)/\psi$, then for any $p \in P$, we have that $d(q, p) \leq (1 + \psi)d(q, P)$.

9.B. (10 PTS.) Assume you are given a set of points P, and a parameter $\delta \in (0, 1)$. Assume that $\text{spread}(P) = O(n^{O(1)})$, where the spread is the ratio between the diameter of P and closest-pair distance in P.

Show how to construct, using the above data-structure, a data-structure of size $\widetilde{O}(n)$, such that given a query point $q \in Z$, one can return in $O(\log n)$ time, a point $p \in P$, such that $d(q, p) \leq (1 + \delta)d(q, P)$. You can safely assume that $\varepsilon > 1/n$, where $n = |P|$.

Show that the total space required by your data-structure is $\widetilde{O}(n)$.

9.C. (20 PTS.) Using 5.B from homework 2, show how to construct a data-structure as above, for the case that the spread of P is unbounded. Show that the space used by your data-structure is $\widetilde{O}(n^2)$.

9.D. (60 PTS.) A natural approach to reduce the space requirement, is to pick a distance r (say, the median edge length in the MST of P), and build a data-structure that answers the ANN query on P, if $\Delta = d(q, P) \in [r/n^2, rn^2]$. If the point fails to be in $\Delta > rn^2$ then the data-structure recursively continues the search on roughly half the points of P (but which ones?). Similarly, if $\Delta < r/n^2$, the search continues recursively on a subset of P that is at most of size $n/2 + 1$ (but much smaller on "average").

Show how to make this scheme work, and prove that it returns the desired ANN, using $\widetilde{O}(n)$ space, answering a ANN query in $O(\log n)$ time.