Tight Lower Bounds for Halfspace Range Searching

Sunil Arya, David M. Mount and Jian Xia

Presented by Josef Cibulka

• Given is a set P of n points in $\mathbb{R}^d$ with weights from a commutative semigroup. The task is to preprocess the set to be able to quickly answer queries asking the sum of elements lying within any halfspace.
• $t(n,m)$ ... time needed to answer a query in the optimal algorithm with $O(m)$ memory in the case of an idempotent semigroup ($\forall x : x + x = x$)
• $t'(n,m)$ ... time needed to answer ... in the case of an integral semigroup ($\forall k \geq 2, \forall x$: the k-fold sum $x + \cdots + x$ is not equal x)
• $\tilde{O}$, $\tilde{\Omega}$ ... ignoring $\log^{O(1)}(n)$–factors
• $\mu(K)$ ... measure of $K \subset \mathbb{R}^d$

Previous bounds. *When $n \leq m \leq n^d$,*

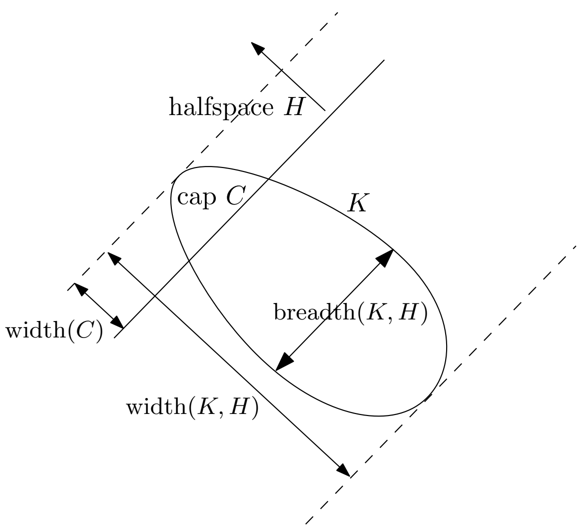
$$\tilde{\Omega}\left(n^{1-(d-1)/(d^2+d)}/m^{1/d}\right) \leq t(n,m) \leq O\left(n^{1-1/(d+1)}/m^{1/(d+1)}\right)$$
$$\tilde{\Omega}\left(n/m^{(d+1)/(d^2+1)}\right) \leq t'(n,m) \leq O\left(n/m^{1/d}\right).$$

The upper bound on $t(n,m)$ assumes that P is uniformly distributed.

Theorem 1. *New lower bounds:*

$$\tilde{\Omega}\left(n^{1-1/(d+1)}/m^{1/(d+1)}\right) \leq t(n,m)$$
$$\tilde{\Omega}\left(n/m^{1/d}\right) \leq t'(n,m)$$

Assumptions.
• *Preprocessing ... determining sums in* generators *(sets G such that $conv(G) \cap P = G$); $\mathcal{G}$ ... set of used generators, $m = |\mathcal{G}|$*
• *Time of query for a halfspace H ... minimum number of generators whose union is $H \cap P$.*
• *In the integral case, the generators are disjoint.*



Choice of P and the ranges.
• *P is inside the unit cube $\mathbb{U} = [0,1]^d$*

- P *is* scattered, *that is, for every* $K \subset \mathbb{U}$*:*
  - $\mu(K) \geq a \log n/n \Rightarrow |K \cap P| \geq (n/a)\mu(K)$
  - $|K \cap P| \geq \log n \Rightarrow \mu(K) \geq |K \cap P|/(na)$
- H_q *… halfspace containing the origin* O *and with boundary passing through* q *and orthogonal to* qO
- *the query range is a random hyperplane* H_q *where* q *is taken from the ball of radius* $1/4$ *centered at* $(1/2, \dots, 1/2)$ *using probability measure* $\int(1/\|q\|^{d-1})dx_1...dx_d$

• A *slab* $S^\Delta(H)$ of width Δ is the set of points of H at distance at most Δ from the bounding hyperplane of H.
• Given n, m and $t = t(n,m)$, let $\Delta_0 = c_0 t \log n/n$. We will consider the slab $S^{\Delta_0}(H)$ (the *region of interest*) of a randomly picked halfspace H and show that many of the generators may be needed to contain all the points of P in the slab without containing anything outside H.

Lemma (Chazelle 1989)**.** *For a convex body* $K \subset \mathbb{U}$*, a random* H *from the above probabilistic space and* $\Delta > 0$*:*

$$\mu(K) Pr[K \subset S^\Delta(H)] = O(\Delta^{d+1}).$$

Idea of the proof in the idempotent case

Lemma (Lemma 4.2)**.** *Consider a convex body* $K \subset \mathbb{R}$ *of surface area* $O(1)$ *and real numbers* $\Delta, v > 0$*. There are* $O(\Delta/v)$ *convex bodies* $K_1, K_2, \dots$ *(*Macbeath regions*) such that for every cap* C *of* K *with width*$(C) \leq \Delta$ *and* $\mu(C) \geq v$ *one of* K_i *satisfies:*
  - $\mu(K_i) \geq \Omega(\mu(C))$ *and*
  - $K_i \subset C$*.*

• The proof of the idempotent case proceeds by considering a generator G lying inside H and containing at least $\log n$ points inside the region of interest R_H (*"interesting" generator*). Then $\mu(\text{conv}(G) \cap R_H)$ is large (since P is scattered), so one of the Macbeath regions of the generator lies within the region of interest (by Lemma 4.2), but this has small probability by the Lemma of Chazelle.

Idea of the proof in the integral case

• A convex body K is α-fat if there are two concentric balls B^- and B^+ such that $B^- \subset K \subset B^+$ and with ratio between the radii of B^+ and B^- at most α.

Lemma (Lemma 4.5)**.** *Consider an* α*-fat compact convex body* $K \subset \mathbb{R}^d$ *and two parameters* $\beta \geq 1$ *and* $\Delta > 0$*. There exists a collection of* $O(\beta \log \alpha)$ *convex bodies* $K_1, K_2, \dots \subset K$ *satisfying the following property: Let* H *be a halfspace, and let* C *be the cap* $K \cap H$*. If width*$(C) \leq \Delta$ *and* $\mu(C) \geq$ *breadth*$(K,H)\Delta/\beta$*, then some* K_i *satisfies:*
  - $\mu(K_i) \geq \Omega(\mu(C))$ *and*
  - $K_i \subset C$*.*

• To be able to apply Lemma 4.5, we say that a generator G is *interesting* if it is lying inside H, contains at least $\log n$ points inside the region of interest R_H and either $G \subset R_H$ or

$$|G \cap R_H| \geq \frac{\text{breadth}(G,H)\Delta_0 n}{c \log(1/\Delta_0)}.$$

It is shown that most points in the region of interest lie in the interesting generators, since otherwise the generators covering the region of interest would have large breaths and would not fit inside $\mathbb{U}$ (since they are disjoint).