18.409 The Behavior of Algorithms in Practice 5/2/2002

Lecture 19

Lecturer: Dan Spielman *Scribe: Matthew Lepinski*

Where We Are in the Proof

We are trying to bound the quantity:

$$\Pr_{b_1,\ldots,b_d}[\text{dist}(0, \partial(\triangle(b_1 \ldots b_d))) < \epsilon]$$

Where the probability is over the distribution with density

$$(\prod_{i=1}^{d} \mu_i(b_i))[0 \in \triangle(b_1 \ldots b_d)]vol(\triangle(b_1 \ldots b_d))$$

Also recall that the μ_i are Gaussian with variance $\sigma^2 \leq 1$ and have centers of norm $\leq \Gamma \leq 1 + 4\sqrt{d \log(n)}$.

Last Time We Proved

$$\Pr_{b_1,\ldots,b_d}[\text{dist}(b_1, \text{dist}(b_2, \ldots, b_d)) < \epsilon] \leq e^2 \left(\frac{d(1+\Gamma)\epsilon}{\sigma^2}\right)$$

Where $\text{dist}(b_2, \ldots, b_d)$ denotes the affine span of $b_2, \ldots, b_d$.

Today We Will Prove

$$\Pr_{b_1,\ldots,b_d}[\text{dist}(0, \text{dist}(b_2, \ldots, b_d)) < \epsilon \text{dist}(b_1, \text{dist}(b_2, \ldots, b_d))] \leq \left(\frac{10d\Gamma^2\epsilon}{\sigma^2}\right) \quad (1)$$

How Smooth are Gaussians

We will now prove a utility lemma which will be useful later.

Lemma 1 *Let μ is a Guassian distribution centered at 0 with variance σ^2. If X and Y are points such that $\|X\| \leq T$ and $\|X - Y\| < \epsilon \leq T$, then*

$$\frac{\mu(Y)}{\mu(X)} \geq e^{\frac{-3\epsilon T}{2\sigma^2}}$$

Proof The worst case is when $X = T$ and $Y = T + \epsilon$. In this case we have:

$$\frac{\mu(Y)}{\mu(X)} = \frac{e^{\frac{-(T+\epsilon^2)}{2\sigma^2}}}{e^{\frac{-T^2}{2sigma^2}}} = e^{\frac{-2T\epsilon-\epsilon^2}{2\sigma^2}}$$

■

Note: If $\epsilon < T/\sigma^2$, this Lemma gives us that

$$\frac{\mu(Y)}{\mu(X)} \geq e^{\frac{-3}{2}}$$

Proof of Today's Bound

Goal: Prove the origin is not much closer to $\text{dist}(b_1, \ldots, b_d)$ than b_1 is.

Note: Since Gaussians are smooth, the chance of any nearby configuration is about the same.

Idea: Fix the shape of the simplex and shift it a little towards b_1. The resulting configuration is just as likely and has the origin farther from $\text{dist}(b_1, \ldots, b_d)$.

Change of Variables

It will be easier to fix the shape of the simplex if we do another change of variables.

Our handle on the simplex will be its center of gravity:

$$x = \frac{1}{d}\sum_{i=1}^{d} b_i$$

The shape of the simplex will be determined by the values:

$$\delta_i = x - b_i$$

For $i = 2 \ldots d$.

Additionally, we set

$$\sum_{i=1}^{d} \delta_i = 0$$

which defines δ_1.

Observe that

$$\triangle(b_1 \dots b_d) = \triangle(x - \delta_1 \dots x - \delta_d) = x + \triangle(-\delta_1 \dots - \delta_d)$$

Therefore,

$$0 \in \triangle(b_1 \dots b_d) \Leftrightarrow x \in \triangle(-\delta_1 \dots - \delta_d)$$

Similarly,

$$\mathrm{dist}(0, \mathrm{dist}(b_2, \dots, b_d) = \mathrm{dist}(0, \mathrm{dist}(\delta_2, \dots, \delta_d)$$

Note: This change of variables is just a linear transformation and so the Jacobian of the transformation is constant.

Defining the Contraction Map

We observe that:

$$\Pr_{b_1,\dots,b_d}[\mathrm{dist}(b_1, \mathrm{dist}(b_2, \dots, b_d)) < \epsilon] \leq \max_{\delta_1,\dots,\delta_d} \Pr_x[\mathrm{dist}(x, \mathrm{dist}(\delta_2, \dots, \delta_d)) < \epsilon \mathrm{dist}(\delta_1, \mathrm{dist}(\delta_2, \dots, \delta_d))]$$

Subject to the condition that $\|\delta_1 - \delta_i\| \leq 2\Gamma$ for all i and where x is has distribution:

$$\nu(x) = (\prod_{i=1}^{d} \mu_i(b_i))[x \in \triangle(\delta_1 \dots \delta_d)]$$

Let S be $\triangle(\delta_1 \dots \delta_d)$. Let S_ϵ be obtained by contracting S at δ_1 by a factor of $(1-\epsilon)$. That is, S_ϵ is the set of points y such that

$$\mathrm{dist}(y, \mathrm{dist}(\delta_2, \dots, \delta_d)) \geq \epsilon \mathrm{dist}(\delta_1, \mathrm{dist}(\delta_2, \dots, \delta_d))$$

We observe that

$$\Pr_x[\mathrm{dist}(x, \mathrm{dist}(\delta_2, \dots, \delta_d)) < \epsilon \mathrm{dist}(\delta_1, \mathrm{dist}(\delta_2, \dots, \delta_d))] = \frac{\nu(S) - \nu(S_\epsilon)}{\nu(S)}$$

To show this quantity is small, we just need to show that the probability measures don't change much.

Let ρ_ϵ be the contraction map specified above.

Bounding $\nu(S_\epsilon)/\nu(S)$

We observe that ρ_ϵ moves points by at most a distance of $\epsilon 2\Gamma$. Also, we know that x is at a distance of at most 3Γ from the center of its distribution.

Therefore, from Lemma 1, we know that under the map, ρ_ϵ, the product of the μ's changes my at most a factor of:

$$e^{\frac{-3-6\epsilon\Gamma^2 d}{2\sigma^2}}$$

Which is at least:

$$\left(1-\frac{\epsilon 9\Gamma^2 d}{\sigma^2}\right)$$

Additionally, are contraction map has a Jacobian of $(1-\epsilon)^d$ at every point.

Note: $(1-\epsilon)^d \leq (1-\epsilon d)$

This implies that

$$\frac{\nu(S_\epsilon)}{\nu(S)} \leq \left(1-\frac{\epsilon 9\Gamma^2 d}{\sigma^2}\right)(1-\epsilon d) \leq \left(1-\frac{10d\epsilon\Gamma^2}{\sigma^2}\right)$$

Therefore,

$$\frac{\nu(S)-\nu(S_\epsilon)}{\nu(S)} \leq \frac{10d\epsilon\Gamma^2}{\sigma^2}$$

Completing the Bound

Using Combination Lemma 2 which we will prove in the next section, we get that:

$$\Pr[\mathrm{dist}(0, \mathrm{dist}(b_2, \ldots, b_d)) < \epsilon] \leq \epsilon \frac{10ed^2\Gamma^2(1+\Gamma)}{\sigma^4}$$

Therefore, since the simplex has d symmetric faces, we get that:

$$\Pr[\mathrm{dist}(0, \partial(b_1, \ldots, b_d)) < \epsilon] \leq \epsilon \frac{10ed^3\Gamma^2(1+\Gamma)}{\sigma^4}$$

The Combination Lemma

Lemma 2 *Let a, b have some distribution such that*

1. $\Pr[f(a) < \epsilon] \leq \alpha\epsilon$

2. $\forall a$, $\Pr[g(a,b) \leq \epsilon] \leq (\beta\epsilon)^2$

Then $\Pr[f(a)g(a,b) < \epsilon] \leq 5\alpha\beta\epsilon$

Note: To really apply this to the previous circumstances, we would need to do this in the δ world where

- $a = \{\delta_1, \dots \delta_2\}$
- $b = x$

Proof

$$\begin{aligned}\Pr[f(a)g(a,b) < \epsilon] \leq \Pr[f(a) < \beta\epsilon] + \sum_{i\geq 1} \Pr[f(a) < \beta\epsilon 2^i \text{ and } g(a,b) < \frac{2^{i+1}}{\beta} \\ \leq \alpha\beta\epsilon + \sum_{i\geq 1} \alpha\beta\epsilon 2^i (2^{-i+1})^2 = \alpha\beta\epsilon(1 + \sum_{i\geq 2} 2^{-i+2})\end{aligned}$$

■

Back to the Original Problem

Recall that the quantity which we were originally interested in was:

$$\Pr_{\omega, r, b_{\pi_1}, \dots, b_{\pi_d}} \left[\frac{\text{dist}(z^{\omega,r}, \partial\triangle(b_{\pi_1}\dots b_{\pi_d})) < \omega, z >}{3\Gamma} < \epsilon \right]$$

Where the distribution of ω, r, and b_i was:

$$\left(\prod_{j \notin \{\pi_1, \dots, \pi_d\}} \int_{a_j} [< \omega, a_j > \leq r] \mu_i(a_j) \right) \prod_{i=1}^{d} \mu_{\pi_i}(a_{\pi_i}) \\ [z^{\omega,r} \in \triangle(b_{\pi_1}\dots b_{\pi_d})] \text{Vol}(\triangle(b_{\pi_1}\dots b_{\pi_d}))$$

We have succeeded in bounding the probability that $z^{\omega,r}$ is close to the boundary of the triangle. Therefore, all that remains is to show it is unlikely that the angle is small. (That is, it is unlikely that $< \omega, z >$ is small).

Idea: Rotate the plane specified by ω and r while preserving the intersection of the plane with the ray z.

Another Change of Variables

Previously, we had been specifying the plane with the variables ω and r. Now we will instead specify the plane by the variables ω and t where tz is the point where the ray z intersects the plane.

Note: $r = t < \omega, z >$

Jacobian $\frac{\partial r}{\partial t} =< \omega, z >$

Also, for convenience we choose $tz = z^{\omega,r}$ to be the origin of the plane containing the b_{π_i}'s.

We will now bound the quantity:

$$\max_{b_{\pi_1},\ldots,b_{\pi_d},t} \Pr_{\omega}[< \omega, r >< \epsilon]$$

Subject to the constraint that $\|b_{\pi_i}\| \leq \Gamma$ and $|t| \leq \Gamma$. Where the distribution for ω is:

$$< \omega, z > \left(\prod_{j \notin \{\pi_1,\ldots,\pi_d\}} \int_{a_j} [< \omega, a_j > \leq t < \omega, z >] \mu_i(a_j) \right) \prod_{i=1}^{d} \mu_{\pi_i}(a_{\pi_i})$$

Observe that in the above expression everything except the term $< \omega, z >$ is like a constant for small changes in ω. We will write this distribution function as $< \omega, z > f(\omega)$.

"Longitude and Latitude"

We will now change to "Longitude and Latitude" where we express the unit vector ω as an angle, θ (latitude) and a point, ϕ on the unit sphere of $d-1$ dimensions.

That is, $\omega \in S^d$ becomes $\theta \in [0, \pi/2]$ and $\phi \in S^{d-1}$. Where $\cos(\theta) =< \omega, z >$.

The Jacobian of this transformation is $[\sin(\theta)]^{d-1}$ which is like a constant for $\theta \to \pi/2$.

Therefore, we get that:

$$\max_{b_{\pi_1},\ldots,b_{\pi_d},t} \Pr_{\omega}[< \omega, r >< \epsilon] \leq \max_{b_{\pi_1},\ldots,b_{\pi_d},t,\phi} \Pr_{\theta}[\cos(\theta) < \epsilon]$$

Where the density is the same as before except $< \omega, z >$ is replaced by $\cos(\theta)$. That is, the density is $f(\theta)\cos(\theta)$.

Note: We could derive a θ_0 such that $\theta_0 = \text{poly}(n, d, 1/\sigma)$ and $0 \leq \theta \leq \theta' \leq \theta_0$

$$\frac{f(\theta)}{f(\theta')} \leq e$$

(This is what was meant earlier by the comment that f is almost constant for small changes in ω)

Therefore, we could show that:

$$\Pr_{\theta}[\cos(\theta) < \epsilon] < \text{poly}(n, d, \frac{1}{\sigma})\epsilon^2$$

Putting it All Together

We now have bounds showing it is unlikely the distance to the boundary is small and it is unlikely that the angle is small. Therefore, we can apply a Combination Lemma to get the following bound:

$$\Pr_{\omega, r, b_{\pi_1}, \ldots, b_{\pi_d}} \left[\frac{\text{dist}(z^{\omega, r}, \partial\triangle(b_{\pi_1} \ldots b_{\pi_d})) < \omega, z >}{3\Gamma} < \epsilon \right] < \epsilon\text{poly}(n, d, \frac{1}{\sigma})$$

This bound implies that the expected size of the shadow of the polytope is polynomial.