18.409 The Behavior of Algorithms in Practice 4/25/2

Lecture 17

Lecturer: Dan Spielman *Scribe: Steve Weis*

Shadow Vertex Rule Review

To briefly review the shadow vertex rule, suppose we have a convex polytope given as the convex hull of a collection of points, an objective function and an initial vector. Moreover, assume that we know the facet of the polytope pierced by the ray through the initial vector. The algorithm continuously modifies the initial vector until it becomes the objective function, all the while tracing the point where the ray through the vector pierces the hull of the polytope. Of course, when one actually implements the algorithm, it takes discrete steps jumping from facet to facet of the convex hull. We need to try and prove a bound on the number of facets this algorithm will crawl over.

In order to show this, we will look at an evenly distributed collection of M rays in the plane spanned by the initial vector and the objective funciton, and count the number of times two adjacent rays pierce different facets. The probability this occurs is upper bounded by $\frac{c}{M}$. To bound the constant c in this probability, we consider the cone of largest angle around a ray that only pierces the facet the ray pierces, and prove that the probability that this angle is less than ϵ is at most $c\epsilon$.

Dan's Favorite LP

Given an objective function c and a convex hull $CH(0, a_1, \ldots, a_n)$, maximize α such that $\alpha \cdot c \in CH$.

Let $Opt \bigtriangleup_z (a_1, \ldots, a_n)$ denote set of indices of the corners of the simplex on the convex hull pierced by the ray through z.

Theorem 1. *Let c and c' be vectors. Let $a_1, \ldots, a_n$ be Gaussian random vectors centered at $\hat{a_1}, \ldots, \hat{a_n}$, which have norm ≤ 1 and variance σ^2. Then:*

$$E[| \cup_\theta Opt \bigtriangleup_{c \sin\theta + c' \cos\theta} (a_1, \ldots, a_n)|] \leq poly(n, d, \frac{1}{\sigma})$$

This expected value is the number of facets which pass through the plane generated by the objective function vector and the initial vector. Consider $(z_1, \ldots, z_m)$ regularly spaced vectors in the plane defined by $span(c, c')$. We need to measure the probability two adjacent rays do not pierce the same simplex, i.e $Pr[Opt\bigtriangleup_{z_i} \neq Opt\bigtriangleup_{z_{i+1}}]$. By calculating this probability we can take the limit as we increase the number of rays to get the expected value from Theorem 1:

$$E[| \cup_\theta \, Opt \triangle_{c \sin\theta + c' \cos\theta} (a_1, \ldots, a_n)|] = \lim_{M\to\infty} \sum_i Pr[Opt\triangle_{z_i} \neq Opt\triangle_{z_{i+1}}]$$

Definition 2.

$$ang(z, \partial \triangle (a_{\pi_1}, \ldots, a_{\pi_d})) = \min_{x\in\partial\triangle(a_{\pi_1},\ldots,a_{\pi_d})} ang(xOz)$$

$$ang_z(a_1, \ldots, a_n) = ang(z, \partial Opt \triangle_z (a_1, \ldots, a_n))$$

This definition refers to the angle of a ray with relation to a point on a simplex. Note that where $ang(xOz)$ is the angle between the rays x and z at the origin. If $Opt\triangle_{z_i} \neq Opt\triangle_{z_{i+1}}$ then $ang_{z_i} \leq \frac{2\pi}{M}$, so: $Pr[Opt\triangle_{z_i} \neq Opt\triangle_{z_{i+1}}] \leq Pr[ang_{z_i}(a_1, \ldots, a_n) \leq \frac{2\pi}{M}]$.

Definition 3. *For any z,* $P_z(\xi) = Pr[ang_z(a_1, \ldots, a_n) \leq \xi]$

Using these definitions, we can bound the expected value from Theorem 1:

$$\lim_{M\to\infty} \sum_i Pr[Opt\triangle_{z_i} \neq Opt\triangle_{z_{i+1}}] \leq \max_z \lim_{\xi\to 0} \frac{2\pi}{\xi} P_z(\xi)$$

We are going to make a very brute force argument about the value of $P_z(\xi)$.

Claim 4. *For any z,*

$$P_z(\xi) = \sum_{\pi_1,\ldots,\pi_d} Pr[opt\triangle_z = \{\pi_1, \ldots, \pi_d\} \bigwedge ang(z, \delta \triangle (a_{\pi_1}, \ldots, a_{\pi_d})) \leq \xi] =$$

$$\sum_{\pi_1,\ldots,\pi_d} Pr[opt\triangle_z = \{\pi_1, \ldots, \pi_d\}] Pr[ang(z, \delta \triangle (a_{\pi_1}, \ldots, a_{\pi_d})) | opt\triangle_z = \{\pi_1, \ldots, \pi_d\}]$$

The second line is derived from the basic law of conditional probability. It will suffice to bound the second term only.

Definition 5. $CH_{\pi_1,\ldots,\pi_d}$ *= the event* $a_{\pi_1}, \ldots, a_{\pi_d}$ *are on the convex hull. That is,* $\exists ||\omega|| = 1, r \geq 0$ *such that* $< \omega, a_{\pi_i} >= r$ *for i=1,...,d;* $\forall j \notin \{\pi_1, \ldots, \pi_d\} < \omega, a_{\pi_j} >\leq r$

Definition 6. $Cross_{z,\pi_1,\ldots,\pi_d}$ *= event that the ray through z crosses* $\triangle(a_{\pi_1}, \ldots, a_{\pi_d})$

Note that $opt\triangle_z = \{\pi_1, \ldots, \pi_d\}$ if and only if $CH_{\pi_1,\ldots,\pi_d}$ and $Cross_{z,\pi_1,\ldots,\pi_d}$.

These definitions suggest a change of variables. Replace $a_{\pi_1}, \ldots, a_{\pi_d}$ with $\omega, r, b_{\pi_1}, \ldots, b_{\pi_d}$ where $a_{\pi_1}, \ldots, a_{\pi_d}$ lie in on the plane specified by ω and r, and $b_{\pi_1}, \ldots, b_{\pi_d}$ indicate where on that plane $a_{\pi_1}, \ldots, a_{\pi_d}$ lie. Thus, $b_{\pi_i} \in \mathbb{R}^{d-1}$ while $a_{\pi_i} \in \mathbb{R}^d$. Note that we originally had d^2 variables. There is a one-to-one correspondence with the new variables since b_π contributes $d(d-1)$ variables, ω contributes $(d-1)$ and r is a single variable. Since we have a change of variables, we need to compute its Jacobian. See web page for this lecture for some references.

Dan's Aggravation

To be more concrete, we need to indicate how $a_{\pi_1}, \ldots, a_{\pi_d}$ are computed from $\omega, r, b_{\pi_1}, \ldots, b_{\pi_d}$. Fix some vector z. Choose a basis for the plane in $\mathbb{R}^d$ orthogonal to z through the origin. The points $b_{\pi_1}, \ldots, b_{\pi_d}$ lie in this plane. Now, for $\omega = z, r \geq 0$ we can use $a_{\pi_i =} b_{\pi_i} + rz$. To handle other values of ω, let T_ω be the orthogonal linear transformation that maps z to ω and is the identity on the orthogonal space. In general, we will apply this as follows:

$$a_{\pi_i} = T_\omega(b_{\pi_i} + rz) = T_\omega(b_{\pi_i}) + rz$$

There is one catch in that this is not well-defined for $\omega = -z$. But, this is a set of measure zero, so it does not matter. Now we can define the derivatives in each variable: The Jacobian of this change of variables is the subject of a theorem of Blaschke: $da_{\pi_1} \cdot \ldots \cdot da_{\pi_d} = Vol(\triangle(b_{\pi_1}, \ldots, b_{\pi_d})) \cdot dr \cdot d\omega \cdot \cdot db_{\pi_1} \cdot \ldots \cdot db_{\pi_d}$. This allows us to define the second term of Claim 4:

$$Pr[ang(z, \delta \triangle (a_{\pi_1}, \ldots, a_{\pi_d}))|opt\triangle_z = \{\pi_1, \ldots, \pi_d\}] =$$
$$\int_{\omega,r} \left(\prod_{j \notin \{\pi_1,\ldots,\pi_d\}} \int_{a_j} [< a_j, \omega > \leq r] \mu_j(a_j) \right) \cdot$$
$$\int_{b_{\pi_1},\ldots,b_{\pi_d}} [Cross_z] \cdot [ang(z, \delta \triangle (a_{\pi_1}, \ldots, a_{\pi_d}) \leq \xi] \cdot \mu(a_{\pi_1}) \cdot \ldots \mu(a_{\pi_d}) \cdot Vol(\triangle(a_{\pi_1}, \ldots, a_{\pi_d}))$$

Let $\nu_{\pi_i}^{\omega,r}(b_{\pi_i}) = \mu_{\pi_i}(T_\omega(b_{\pi_i}) + r\omega)$. We want to understand $ang(z, \delta \triangle (a_{\pi_1}, \ldots, a_{\pi_d})$, which is the angle of incidence between the plane and a vector. Let $z_{\omega,r}$ and θ be the point and angle where z intersects the plane. Let x be a boundary point of $\triangle(a_{\pi_1}, \ldots, a_{\pi_d})$ and α be the angle $ang(x0z)$. Let $T = dist(O, z^{\omega,r})$ and $\delta = dist(x, z_{\omega,r})$. From all this, we can arrive at the following bound: $\alpha \approx \tan(\alpha) = \frac{\delta \sin\theta}{T - \delta\cos\theta} \geq \frac{\delta \sin\theta}{2(1+4\sqrt{d \log n})}$. This can replace the $[ang \ldots]$ term in the above integral.

Claim 7. $Pr[\exists i ||a_i|| > 4\sqrt{d\log(n)}] \leq e^{-4d\log n}$

There are at most $\binom{n}{d}$ facets and $\binom{n}{d}e^{-4d\log n} \leq e^{-3d\log n}$. So, assuming $\max ||a_i|| \leq \sqrt{(}d\log(n))$, our estimate is off by at most 1.