18.419 Random Walks and Polynomial-Time Algorithms 2-12-2002

Lecture 3: Convex Optimzation II

Lecturer: Santosh Vempala *Scribe: Chris Peikert*

Last time, we reduced convex optimization to the following problem:

- **Input:** convex body K (specified by a separation oracle), R (the side of an origin-centered cube containing K), r (the side of a cube contained in K, if $K \neq \emptyset$)
- **Output:** a point $x \in K$, or the declaration "K is empty"

We then gave an algorithm to solve the problem (note m will be defined later, during the analysis):

1. Let P be the origin-centered cube of side R; $z \leftarrow 0$
2. Ask the separation oracle if $z \in K$. If yes, return z. Otherwise, the oracle returns a, b such that $K \subseteq \{x \ : \ a^T x \leq b\}$.
3. Let $P \leftarrow P \cap \{x \ : \ a^T x \leq a^T z\}$
4. Choose random $Y^1, Y^2, \ldots, Y^m$ from P. Let $z \leftarrow \frac{1}{m} \sum_{i=1}^{m} Y^i$. Goto step 2.

We proved last lecture that if z is the centroid of P, then each side of the separating cut contains at least $1/e$ of P's volume, so $O(n \log(R/r))$ iterations are sufficient for correctness of the algorithm. For our randomized algorithm, we need to consider the probability that any cut through z leaves a constant fraction of P's volume on each side.

As an illustration, assume we are sampling a single point x from a solid ball $B_n(R)$ of dimension n and radius R, centered at the origin. Then the expected length of x from the center is computed by integrating over shells:

$$E[\|x\|] = \frac{1}{\text{Vol}(B_n(R))} \int_0^R g(n) r^{n-1} \, r \, dr$$

and

$$\text{Vol}(B_n(R)) = \int_0^R g(n) r^{n-1} \, dr$$

which yields $E[|x|] = Rn/(n+1)$. If we take a plane at this distance from the center, the fraction of the ball's volume on one side of the cut is exponentially small in n.

The question is now: how close must a point be to the center in order to have a constant fraction of the ball's volume on each side? If the plane is at distance T from the center, the radius of its cross-section is $\sqrt{R^2 - T^2}$, and the volume of the smaller part is:

$$\begin{aligned} \frac{\mathrm{Vol}(B_n(R))}{2} - \int_{t=0}^{T} g(n)\left(\sqrt{R^2 - t^2}\right)^{n-1} dt &\geq \frac{R \cdot \mathrm{Vol}(B_{n-1}(R))}{2\sqrt{n}} - T \cdot \mathrm{Vol}(B_{n-1}(R)) \\ &= \mathrm{Vol}(B_{n-1}(R))\left(\frac{R}{2\sqrt{n}} - T\right) \end{aligned}$$

where we've used the fact that $\mathrm{Vol}(B_n(R)) = \frac{2\pi^{n/2}R^n}{n\Gamma(n/2)}$, where $\Gamma(n) = (n-1)!$ for integer n. So if $T \leq \frac{R}{4\sqrt{n}}$, then the volume of the smaller part is at least $1/4$ of the total volume. If we then only sample one random point from the sphere per iteration, we expect $\Omega(n^{n/2})$ trials to reduce the volume by a constant fraction!

Now, as in the algorithm, say we choose $Y^1, \ldots, Y^m$ and define $Y = \frac{1}{m}\sum Y^i$. Wlog, we can say $E[Y] = E[Y^i] = 0$ by centering the ball at the origin. A calculation reveals that $E[|Y^i|^2] = R^2 n/(n+2)$, and because the samples are independent, the variance of Y is $E[Y^2] = \frac{1}{m}E[|Y^i|^2] = \frac{nR^2}{(n+2)m}$. If we set $m = 32n$, the variance is $\frac{R^2}{32(n+2)}$. From above, in order to cut at least a constant fraction of volume, we need $|Y|^2 \leq R^2/16n$. Then by Chebychev's inequality, we get $\Pr[|Y|^2 \geq R^2/16n] \leq \frac{16n}{R^2}E[|Y|^2] \leq 1/2$. Therefore only a constant number of iterations are necessary to cut a constant fraction of volume, with high probability.

More generally, this result holds for any ellipsoid, because an ellipsoid is just a ball under a linear transformation, and relative volumes are preserved under linear transformations (volumes are magnified by the determinant of the transformation matrix).

Now we deal with the general case when K is an arbitrary convex set. First we need a definition:

Definition 1 (Isotropic position). *A body K is in* isotropic position *if:*

1. *Its centroid is at the origin, and*
2. *For all v with $|v| = 1$, and for uniformly random x from K, $E[(v^T x)^2] = 1$.*

It will be useful to re-characterize isotropic position in other terms, with the following lemma.

Lemma 2. *K is in isotropic position iff $E[x_i x_j] = 1$ when $i = j$, and $= 0$ otherwise; i.e. $E[xx^T] = I$.*

Proof: Say $E[xx^T] = I$, and $|v| = 1$. Then $E[v^T xx^T v] = v^T E[xx^T] v = |v|^2 = 1$ by linearity of expectation.

For the other direction, say K is in isotropic position and pick pick $v = e_i$ (the unit vector with 1 as its ith entry). Then $1 = E[(v^T x)^2] = E[x_i^2]$. Next, pick $v = \frac{1}{\sqrt{2}}(e_i + e_j)$. Then $1 = E[(v^T x)^2] = E[x_i^2/2 + x_j^2/2 + x_i x_j] = 1 + E[x_i x_j] \Rightarrow E[x_i x_j] = 0$, as desired. □

Now assume that K is not in isotropic position. Then $E_K[xx^T] = M$, which is semidefinite, because for all $y = cv$ with $|v| = 1$, $y^T My = E[(y^T x)^2] = c^2 E[(v^T x)^2] > 0$, since K is full-dimensional. Therefore $M = B^2$ for some symmetric matrix B. Let's now apply the linear transformation $x \to B^{-1}x$ to the space, and define $K' = \{B^{-1}x \ : \ x \in K\}$. Then for random $y \in K'$, $E_{K'}[yy^T] = E_K[B^{-1}xx^T B^{-1}] = B^{-1}E_K[xx^T]B^{-1} = I$. Then K' is in isotropic position, so from now on we will assume wlog that K is in isotropic position because ratios of volumes are preserved under linear transformations.

We will complete the analysis of the algorithm with the following theorem.

Theorem 3. *For an isotropic convex body K, any cut at a distance t from the centroid has at least $\frac{1}{e} - t$ of its volume on each side of the cut.*

We will prove this theorem a bit later. For our purposes, we will need $t < 1/e$. In fact, $E[t^2] = E[|Y|^2] = \frac{1}{m}E[|Y^i|^2] = n/m = 1/32$, by linearity of expectation and the definition of isotropic position. Again, by Chebychev's inequality, we conclude that the algorithm terminates in $O(n \log(R/r))$ iterations in expectation.

To prove Theorem 3, we will need the following lemma:

Lemma 4. *Let y be the variable of distance along direction a, and let K be a convex body. Define $f(y)$ to be the $(n-1)$-dimensional volume of the cross-section at distance y, divided by Vol(K) (so $\int f(y)\,dy = 1$). If $\int y^2 f(y)\,dy = 1$ (i.e., K's second moment is 1) and $\int yf(y)\,dy = 0$, then $\max_y f(y) < 1$.*

We will prove the lemma after applying it. Note that an isotropic body fits the conditions of the lemma.

Proof of Theorem 3: In the language of Lemma 4, let a be normal to, and in the direction of, the plane of the cut. Then the volume ratio is at least $1/e - \int_0^t f(y)\,dy \geq 1/e - [\max_y f(y)] \int_0^t dy \geq 1/e - t$, as desired. □

Proof of Lemma 4: we proceed in a way much like last lecture's proof of constant-ratio cuts. First assume that K is the body which has the *largest* $\max_y f(y)$. Now symmetrize around a, i.e. each cross section of K normal to a is replaced by an $(n-1)$-dimensional ball having the same volume as the cross-section. As we have already seen, the body remains convex, $\max_y f(y)$ remains the same, as do the first and second moments. Assume now wlog that $f(y)$ is maximal at $y^* \geq 0$, and examine sections K_1, K_2, K_3 given by $y < 0$, $0 < y < y^*$, and $y^* < y$, respectively. Define the moment about the centroid $\overline{y}$ to be $I(K) = \int (y - \overline{y})^2 f(y)\, dy$. Now replace K_1 by a cone of the same volume, with same cross-sectional area at $y = 0$ and base at some $y \leq 0$. Replace K_2 by a truncated cone (possibly having smaller volume), retaining cross-sectional areas at $y = 0$ and $y = y^*$. Replace K_3 with a cone of same volume plus whatever was lost from K_2, with same cross-sectional area at $y = y^*$. In all of these transformations, mass moves away from the centroid, so the new moment $I(K')$ cannot be smaller. If it is larger, we can "squeeze" the body inward along a and expand it orthogonal to a (increasing its cross-sectional areas) until its moment is again 1. But this contradicts the assumption that we started with the body of largest $\max_y f(y)$. Therefore the above operations do not move any mass.

We now describe the body having the largest $\max_y f(y)$. By the above, its ends must consist of cones and its middle of a truncated cone. A cone with base area equal to the cross-sectional area at y^* has a larger moment of inertia than K'. Finally, a cone with $\int y^2 f(y)\, dy = 1$ and $\int y f(y)\, dy = 0$ has $\max_y f(y) = \frac{n}{n+1}\sqrt{\frac{n}{n+2}} < 1$, and we are done. $\square$