Handout 4, the Ellipsoid algorithm

January 11, 2023

The Ellipsoid algorithm was developed by (formerly) Soviet mathematicians (Shor (1970), Yudin and Nemirovskii (1975)). Khachian (1979) proved that it provides a polynomial time algorithm for linear programming. The average behavior of the Ellipsoid algorithm is too slow, making it not competitive with the simplex algorithm. However, the theoretical implications of the algorithm are very important, in particular, providing the first proof that linear programming (and a host of other problems) are in P.

Consider a generalization of the “20 questions” game, from one dimension to many dimensions. The input to the game is an n dimensional ball B of radius $R > 1$. In it, an adversary hides a unit ball U. The purpose of the algorithm is to find a point in U. The game proceeds in rounds. In each round, the algorithm specifies one point $p \in B$. If p happens to be in U, the game ends and the algorithm wins. If not, the adversary provides a hyperplane that separates between p and U (such an adversary is referred to as a *separation oracle*), and the game proceeds to the next round. There are algorithms that win the game within $O(n \log R)$ rounds, but their complexity per round is rather high. The Ellipsoid algorithm wins the game within $O(n^2 \log R)$ rounds. It is a polynomial time algorithm, though not *strongly polynomial time*.

Let Q be an n by n nonsingular real matrix and $t \in R^n$. The mapping $T(x) = Qx + t$ is called an *affine transformation*. A *unit ball* $S(0,1)$ in R^n is the set $\{x | x^T x \leq 1\}$. An *ellipsoid* is the image of a unit ball under an affine transformation. Observe that $y = Qx + t$ implies $x = Q^{-1}(y - t)$. Hence using the notation $B = QQ^T$, an ellipsoid is

$$T(S(0,1)) = \{y | (Q^{-1}(y-t))^T Q^{-1}(y-t) \leq 1\} = \{y | ((y-t)^T B^{-1}(y-t) \leq 1\}.$$

The matrix B is *positive definite* meaning that it is real and symmetric, and satisfies the following conditions (all of which are equivalent):

- $x^T Bx > 0$ for all nonzero $x \in R^n$.
- all its eigenvalues are real and positive.
- there exists a matrix Q with linearly independent rows such that $B = QQ^T$. (One possible choice for Q is to have column i equal to $\sqrt{\lambda_i} v_i$, where v_i is the ith eigenvector of B, and λ_i is its eigenvalue.)

The eigenvectors of B are the principle axes of the ellipsoid, the square roots of the eigenvalues are their lengths, and the square root of the determinant gives the volume (scaled by the volume of the unit ball).

In the ellipsoid algorithm we construct a sequence of ellipsoids $E_k = (B_k, t_k)$. If t_k violates the constraint $a_i^T x \leq b_i$ then we take E_{k+1} to be an ellipsoid that contains $\frac{1}{2}E_k = \{y \in E_k : a_i^T y \leq a_i^T t_k\}$, for which there are the following formulas:

$$t_{k+1} = t_k - \frac{1}{n+1}\frac{B_k a_i}{\sqrt{a_i^T B_k a_i}} \qquad B_{k+1} = \frac{n^2}{n^2-1}(B_k - \frac{2}{n+1}\frac{B_k a_i a_i^T B_k}{a_i^T B_k a_i})$$

It can be shown that $\mathrm{vol}(E_{k+1}) < e^{-1/2(n+1)}\mathrm{vol}(E_k)$. The proofs start with the simplest ellipsoid, the unit ball, and then use linear transformations.

Positive semidefinite matrices form a convex set, and the ellipsoid algorithm can be used to optimize over them (*positive semidefinite programming*) up to arbitrary precision. As an example, we consider the problem of embedding a finite metric space in Euclidean space with minimum distortion.

Homework: Hand in by January 25 (though note that there will be no class on January 25).

1. Prove that for any point in a triangle, there is a line that passes through the point and separates the triangle into two regions, one of which has at least a 5/9-fraction of the area of the triangle. (Hint: consider what happens if the point is the center of gravity of the triangle and the line is parallel to an edge of the triangle.)

2. Prove that the set of n by n positive definite matrices is a convex set. (Namely, for every $0 \leq \alpha \leq 1$ and two positive definite matrices A and B of order n, the matrix $C = \alpha A + (1-\alpha)B$ is positive definite.)