Mathematical Programming II	**Lecture 7**
ORIE 6310 Spring 2014	**February 13, 2014**
Scribe: Zhijie(Sasha) Dong	

The Complexity of Pivoting Algorithms

The example last time, with similar examples for most variants of the simplex method, gives a discouraging view of the efficiency of pivoting algorithms.

These examples show bad worst-case behavior of such methods, but they seem to work well in practice, typically taking a **linear** (in the dimensions) number of steps. Can we explain this? We are going to discuss this question from two points of view:

(i) The **diameter** of polyhedra.

(See the books by Ziegler and Kalai-Ziegler on reserve.)

(ii) Polynomial **expected** behavior of a pivoting algorithm.

(See Prof. Todd's paper "Polynomial expected behavior ..." on reserve and on the homepage.)

Definition 1 *A* ***polytope*** *is the convex hull of a finite number of points:* $P = \left\{V\lambda : e^T\lambda = 1, \lambda \geq 0\right\}$, *for some* $V \in \mathbb{R}^{d\times k}$*, where* d *is the dimension.*

A ***polyhedron*** *is the intersection of a finite number of halfspaces:* $P = \left\{x \in \mathbb{R}^d : Ax \leq b\right\}$, *for some* $A \in \mathbb{R}^{n\times d}, b \in \mathbb{R}^n$.

Remark 1 *A* ***bounded*** *polyhedron is a polytope and vice versa.*

Definition 2 *Given a* d*-dimensional polyhedron* P, $v \in P$ *is a* ***vertex*** *of* P *if for some* $c \in \mathbb{R}^d$, $\operatorname{argmin}\left\{c^Tx : x \in P\right\} = \{v\}$. *For two vertices* v *and* w, v *is* ***adjacent*** *to* w *in* P *and* $[v,w] = \left\{(1-\lambda)v + \lambda w : 0 \leq \lambda \leq 1\right\}$ *is an* ***edge*** *of* P *if there is some* $c \in \mathbb{R}^d$ *with* $\operatorname{argmin}\left\{c^Tx : x \in P\right\} = [v,w]$.

Definition 3 *A* ***path*** *in* P *from* v *to* w *of* ***length*** k *is a sequence* $v = v_0, v_1, ..., v_k = w$*, with* v_{i-1} *and* v_i *adjacent for* $i = 1, ..., k$. *The* ***distance*** *from* v *to* w *in* P *is the length of the shortest such path,* $d_P(v,w)$.

Definition 4 *The* ***diameter*** *of* P *is the largest such distance,*
$\delta(P) := max\left\{d_P(v,w)\ v \text{ and } w \text{ are vertices of } P\right\}$

P *is* ***neighborly*** *(or 1-neighborly) if every two vertices are adjacent, so* $\delta(P) = 1$.

The diameter of a polyhedron being "small" implies that some simplex variant may be "efficient"; the diameter of a polyhedron being "large" implies that **all** simplex variants must be "bad".

Clearly, $\delta([0,1]^d) = d$, $\delta(\operatorname{conv}\left\{\pm e_i, i = 1, ..., n\right\}) = 2$, and $\delta(\left\{x \in \mathbb{R}^d : x \geq 0, e^Tx \leq 1\right\}) = 1$. The first is the cube, the second the cross-polytope (octahedron in dimension 3) and the third the simplex. One might guess the simplex is the only neighborly polytope.

Definition 5 *Choose any* $n \geq d+1$, *and* n *points on the moment curve* $\{(t; t^2; ...; t^d), t \in \mathbb{R}\}$ *corresponding to* $t_1 < t_2 < ... < t_n$. *Their convex hull is called a* ***cyclic polytope***, P.

Lemma 1 *The point $(t; t^2; ...; t^d)$ lies on the hyperplane $H = \{x \in \mathbb{R}^d : a_1x_1 + ... + a_dx_d = -a_0\}$* ***if and only if*** *$p(t) := a_0 + a_1t + ... + a_dt^d$ vanishes, and on its positive side $\{x \in \mathbb{R}^d : a_1x_1 + ... + a_dx_d > -a_0\}$ iff $p(t) > 0$, and on its negative side $\{x \in \mathbb{R} : a_1x_1 + ... + a_dx_d < -a_0\}$ iff $p(t) < 0$.*

Remark 2 *We say H* ***corresponds*** *to P.*

Theorem 1 *For any $d \geq 4$, there are neighborly polytopes with* ***arbitrarily many*** *vertices.*

Proof: Using Lemma 1 and the definitions, we first show that each point $v^i = (t_i; t_i^2; ...; t_i^d)$ is a vertex of P. Consider $p_i(t) = (t - t_i)^2$, which is equal to 0 for $t = t_i$, and positive for $t \neq t_i$. This gives a corresponding hyperplane, and hence an objective function, with c^i showing v^i a vertex. Next, each pair of points (v^i, v^j) defines an edge, since $p_{ij}(t) := (t - t_i)^2(t - t_j)^2$ is equal to 0 for $t = t_i$ or t_j, and positive otherwise. This gives us the desired objective function showing $[v_i, v_j]$ an edge: any convex combination x of the vertices must have $p_{ij}(x)$ positive unless the weights on all vertices except v^i and v^j are zero. Note that we can only use this in $\mathbb{R}^d$ for $d \geq 4$, since p_{ij} has degree 4.

Note: The cyclic polytope has every vertex with $n - 1$ incident edges, compared to d for a simple polyhedron (i.e., nondegenerate in the LP sense), so it is massively degenerate, hence maybe not so good for the simplex method, even if it is neighborly.

The cyclic polytope with n vertices in $\mathbb{R}^d$ $(n \geq d + 1)$:

If $d \geq 4$, it has the maximum number of edges $\binom{n}{2}$;

If $d \geq 6$, it has the maximum number of 2-dimensional faces $\binom{n}{3}$ (since we can use $(t - t_i)^2(t - t_j)^2(t - t_k)^2$ for any i, j, k);

......

In general, it has the maximum number of $\lfloor \frac{d}{2} \rfloor - 1$-dimensional faces $\binom{n}{\lfloor \frac{d}{2} \rfloor}$. In fact,

Theorem 2 *(Upper Bound Theorem of McMullen) The cyclic polytope has the maximum number of* ***any*** *dimensional face, and it has $\binom{n-\lfloor \frac{d+1}{2} \rfloor}{n-d} + \binom{n-\lfloor \frac{d+2}{2} \rfloor}{n-d}$ $(d-1)$-dimensional faces (facets).*

Consider a bounded polyhedron with 0 in its interior, $P := \{x : a_i{}^T x \leq 1, \text{for } \ i = 1, ..., n\}$. Consider $P^* := \{y : x^T y \leq 1, \text{all } \ x \in P\}$, the **polar** of P. Then $P^* = \text{conv}(\{a_1, ..., a_n\})$ Also, $P^{**} = P$. The k-dimensional faces of $P \Leftrightarrow (d - k + 1)$-dimensional faces of P^*.

So the polar of a cyclic polytope (translated so that 0 lies in its interior) gives a bounded polyhedron defined by n inequalities with $\binom{n-\lfloor \frac{d+1}{2} \rfloor}{n-d} + \binom{n-\lfloor \frac{d+2}{2} \rfloor}{n-d}$ (the maximum number of) vertices. This is super-exponential.

Let $\Delta(d, n) := \max \{\delta(P) : P \text{ is a } d\text{-dimensional polyhedron with } n \text{ facets}\}$,

$\Delta_b(d, n) := \max \{\delta(P) : P \text{ is a bounded } d\text{-dimensional polyhedron with } n \text{ facets}\}$.

Hirsch conjecture, 1957: $\Delta(d, n) \leq n - d$ *(shown false in 1967 by Klee and Walkup).*

Variant: open till 2010: $\Delta_b(d, n) \leq n - d$, which indicates that we can move from any basic feasible solution to any another in m $(m := n - d)$ steps.

Santos, 2010: Constructed a counterexample P: $\delta(P) = 44$, where P is a 43-dimensional polytope with 86 facets.
Matschke, Santos and Weibel, 2012: A smaller counterexample: $\delta(P') = 21$, where P' is 20-dimensional polytope with 40 facets.
So, $\Delta_b(d, n) \geq \frac{21}{20}(n - d)$, and Klee and Walkup showed $\Delta(d, n) \geq \frac{5}{4}(n - d)$. But while the conjectures have been disproved, the big open question remains: are Δ and Δ_b polynomial in n and d, or even linear?

Reference: paper by Ziegler, "Who solved the Hirsch conjecture?", on the homepage.