

Abstract

There are algorithms with exponential worst-case complexity but which work very well in practical applications. Probably the most famous representative of such an algorithm is the Simplex Method for Linear Programming. This gap between the worst-case complexity and the practically observed performance motivates doing a so called probabilistic analysis. Most important to mention in our case are the Average-Case-Analysis done by Borgwardt and the Smoothed Analysis done by Spielman and Teng. Both analyses base on the shadow-vertex pivot rule for the Simplex Method. Differences in both approaches can be found in the used techniques to derive upper bounds for the expectation values determining the complexity in each case.

The Complexity of Algorithms

We want to state three different complexity types for an algorithm. Therefore, let $C_{\mathcal{A}}$ be a complexity measure of an algorithm \mathcal{A} on input x . Ω shall be the domain of all possible inputs for \mathcal{A} and Ω_n the domain of all inputs of length n .

Worst-Case Complexity:

The algorithm \mathcal{A} has worst-case C -complexity $f(n)$, if $\max_{x \in \Omega_n} \{C_{\mathcal{A}}(x)\} = f(n)$ for a function f from \mathbb{N} into \mathbb{R}^+ .

Average-Case Complexity:

For a family μ of distributions μ_n on Ω_n , we say that \mathcal{A} has average-case C -complexity $g(n)$ under μ , if $\mathbb{E}_{\mu_n} [C_{\mathcal{A}}(x)] = g(n)$ where g is a function from \mathbb{N} into \mathbb{R}^+ .

Smoothed Complexity:

\mathcal{A} has smoothed C -complexity $h(n, \sigma)$, if $\max_{x \in \Omega_n} \{\mathbb{E}_g [C_{\mathcal{A}}(x + (\sigma \|x\|_?)g)]\} = h(n, \sigma)$. Here, $h(n, \sigma)$ is a function from $\mathbb{N} \times \mathbb{R}^+$ into \mathbb{R}^+ and $(\sigma \|x\|_?)g$ is a vector of Gaussian random variables of mean 0 and standard deviation $\sigma \|x\|_?$ where $\|x\|_?$ is a measure of the magnitude of x (e.g. the euclidian norm or the maximum norm).

The Shadow-Vertex Algorithm

At the moment we want to deal with problems of the form

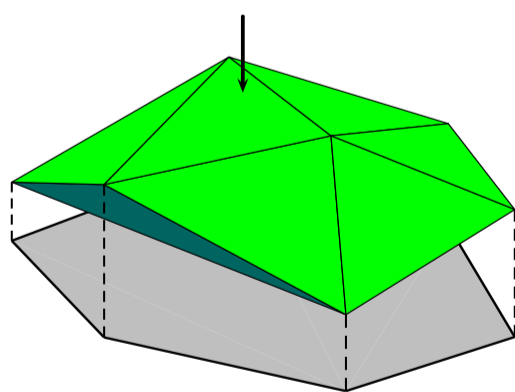
$$\begin{aligned} & \text{maximize } \langle v, x \rangle \\ & \text{subject to } x \in X := \{x \mid \langle a_1, x \rangle \leq 1, \dots, \langle a_m, x \rangle \leq 1\} \end{aligned}$$

where $v, x, a_1, \dots, a_m \in \mathbb{R}^n$ and $m > n$. We want to assume that a vertex x_s of X has already been given so that we can start with *Phase 2* of our Simplex Method right away. This vertex shall be optimal with respect to the problem maximize $\langle u, x \rangle$ for a vector $u \in \mathbb{R}^n \setminus \{0\}$.

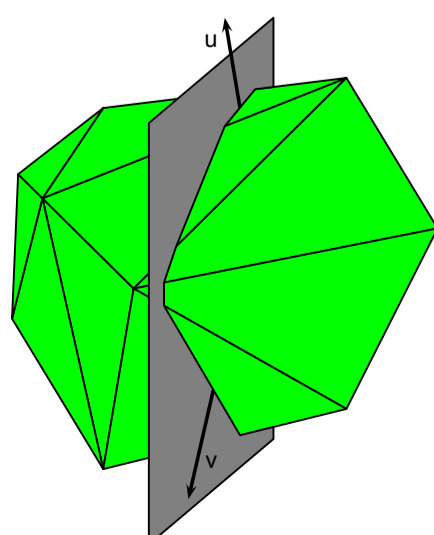
Assumption of Nondegeneracy:

Each n -element subset of $\{a_1, \dots, a_m, u, v\}$ is linearly independent and each subset of $n+1$ elements out of $\{a_1, \dots, a_m\}$ is in general position.

The primal shadow-vertex algorithm constructs a simplex path from x_s to a vertex x_z using all vertices \tilde{x} of X for which there exists a $w \in \text{cone}(u, v)$ such that \tilde{x} is optimal with respect to maximize $\langle w, x \rangle$. x_z is either the optimal solution or the last vertex before an unbounded ray improving the objective function. We can bound the number of vertices on this simplex path by counting the number S of vertices of X which are optimal for any objective function $t \in \text{span}(u, v)$. It can be proven that S is equal to the number of vertices of the two-dimensional polygon which results from the orthogonal projection of X on $\text{span}(u, v)$.



Now we switch to the dual polyhedron Y (with respect to X) as it is easier to calculate S in this dual perspective. Y is defined as the set $\{y \mid \langle x, y \rangle \leq 1 \forall x \in X\}$ and it is possible to prove that we can also express it as $\text{conv}(0, a_1, \dots, a_m)$. Exploiting our nondegeneracy condition and the properties of Y , we can deduce that S is equal to the number of n -element subsets $\Delta = \{\Delta^1, \dots, \Delta^n\}$ of $\{1, \dots, m\}$ for which $\text{conv}(a_{\Delta^1}, \dots, a_{\Delta^n})$ is a $(n-1)$ -dimensional side-simplex of Y and intersected by $\text{span}(u, v)$.



The Average-Case-Analysis

Our stochastic model:

The random vectors a_1, \dots, a_m, v shall be distributed **independently, identically** and **symmetrically under rotations** on $\mathbb{R}^n \setminus \{0\}$.

Based on the dual shadow-vertex algorithm, we can obtain the following integral formula for the expectation value of S :

$$\begin{aligned} \mathbb{E}_{m,n}[S] &= (\text{expected number of } \Delta\text{'s, such that } \text{conv}(a_{\Delta^1}, \dots, a_{\Delta^n}) \text{ is a} \\ & \quad \text{facet of } Y \text{ and intersected by } \text{span}(u, v)) \\ &= (\text{number of candidates } \Delta) \cdot (\text{probability, that the candidate} \\ & \quad \Delta' = \{1, \dots, n\} \text{ satisfies both conditions}) \\ &= \binom{m}{n} \cdot \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} \underbrace{G(h(a_1, \dots, a_n))^{m-n} \cdot n \cdot W(a_1, \dots, a_{n-1})}_{(*)} \\ & \quad f(a_1) \dots f(a_n) da_1 \dots da_n \end{aligned}$$

In the integral formula, $G(h(a_1, \dots, a_n))$ stands for the probability that another random vector a_j lies on the same side of the hyperplane $\text{aff}(a_1, \dots, a_n)$ than the origin. Moreover, $n \cdot W(a_1, \dots, a_{n-1})$ indicates the probability that $\text{conv}(a_1, \dots, a_n)$ is intersected by $\text{span}(u, v)$. Thus $(*)$ indicates the probability that $\text{conv}(a_1, \dots, a_n)$ is a facet of Y and intersected by $\text{span}(u, v)$. Estimating this integral formula delivers $C \cdot m^{\frac{1}{n-1}} \cdot m^2$ as an upper bound where C is a constant independent of m and n . Based on this we can obtain the following result for the analyzed algorithm:

Theorem:

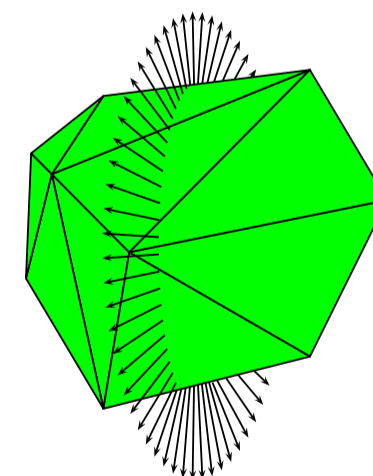
For all distributions according to our stochastic model, the Dimension-by-Dimension Shadow-Vertex Simplex Method does not require more than $C \cdot m^{\frac{1}{n-1}} \cdot n^3$ pivot steps on the average.

The Smoothed Analysis

Our model of perturbation:

The vectors a_1, \dots, a_m are Gaussian random vectors of standard deviation σ , centered at points of norm at most 1. Moreover, u and v are fixed vectors and thus independent of the a_i 's.

Instead of determining the number of facets of Y which are intersected by $\text{span}(u, v)$, we calculate the number of facets intersected by a finite number of vectors from $\text{span}(u, v)$.



Let q_i be the i -th of these all in all l vectors and let E_i be a random variable indicating whether the i -th and the $(i+1)$ -th vector intersect different facets of Y . Then, for l sufficiently big, $S = \sum_{i=1}^l E_i$ holds and thus $\mathbb{E}_{a_1, \dots, a_m} [S] = \mathbb{E}_{a_1, \dots, a_m} \left[\sum_{i=1}^l E_i \right]$. Excessively exploiting the above mentioned model of perturbation (together with the discretization), we can deduce that $\mathbb{E}_{a_1, \dots, a_m} [S] \leq \frac{58,888,678 m^3}{\min(\sigma, \frac{1}{3\sqrt{n \ln m}})^6}$. This is the basis for the further analysis which delivers the following result:

Theorem:

For an arbitrary linear program with $n > 3$ variables and $m > n$ constraints, the expected number of pivot steps of the analyzed two-phase Shadow-Vertex Simplex Method for the smoothed program (perturbed according to the definition of smoothed complexity) is at most a polynomial $\mathcal{P}(m, n, \sigma^{-1})$.

References

1. Karl Heinz Borgwardt. *The Simplex Method: A Probabilistic Analysis*. Number 1 in Algorithms and Combinatorics. Springer-Verlag, Basel, 1987.
2. Daniel Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, Vol 51 (3):pp. 385 - 463, 2004.