11. Rank lower bound for knapsack.

Lecturer: Massimo Lauria

http://www.csc.kth.se/~lauria/sos14/

Disclaimer: this lecture note has not yet been reviewed by the main lecturer. It is released as it is for the convenience of the students.

In this lecture we show that the Knapsack problem requires large Positivstellensatz calculus $\mathrm{PC}_>$ degree to refute, where the Knapsack problem states that a sum of n integer variables is equal to some number r. Formally, we encode the Knapsack problem using the following equations:

$$\begin{aligned} f: &\quad \sum_{i=1}^{n} x_i - r = 0 && \text{where } r \in \mathbb{R} \\ f_i: &\quad x_i^2 - x_i = 0 && \text{for every } i \in \{1,\dots,n\} \end{aligned}$$

Constraints f_i enforce that variables take 0-1 values. Hence, it follows that the constraints are satisfiable if r is an integer between 0 and n, and unsatisfiable if r is outside of that range or non-integral. The latter case of non-integral r is the one we focus on in this lecture. The lecture is based on the paper by Grigoriev, which builds on a result by Impagliazzo, Pudlák, and Sgall [1]. Formally, we show the following theorem.

[1] Dima Grigoriev. Complexity of positivstellensatz proofs for the knapsack. *computational complexity*, 10(2):139–154, 2001; and Russell Impagliazzo, Pavel Pudlák, and Jiri Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999

Theorem 1. *Let r be a real number such that $k < r < n-k$. Then we have that if $0 \le k \le \lceil \frac{n}{4} \rceil - 2$, the degree of $\mathrm{PC}_>$ refutation is lower bounded by $2k+2$. If $k > \lceil \frac{n}{4} \rceil - 2$, then the $\mathrm{PC}_>$ degree is at least $\lceil \frac{n}{2} \rceil + 2$.*

Note that if r is integer, the lower bound is trivial as the constraints are satisfiable and there are no refutations. In the case when $r < 0$ or $r > n$ there are simple constant degree refutations. In general the equational part is sufficient to have $n/2$ degree upper bound for any value of r, by adapting Theorem 4.2 in [2].

[2] Russell Impagliazzo, Pavel Pudlák, and Jiri Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999

As the initial constraints are just equations, we can concentrate on the part of Positivstellensatz calculus which only deals with equations. That is, we concentrate on the proof system that has the following inference rules

$$\textit{linear combination} \quad \frac{p \qquad q}{\alpha p + \beta q}$$

$$\textit{variable multiplication} \quad \frac{p}{x_i p}$$

where p and q are previously derived polynomials or original constraints, and $\alpha, \beta \in \mathbb{R}$. We can derive a polynomial p if

$$p = p' + \sum_i h_i^2 \tag{1}$$

where p' has been inferred using the inference rules and h_i's are arbitrary polynomials. A refutation of initial constraints is the derivation of $p = -1$.

We will show that for the Knapsack problem this full generality is not needed and that the *one-shot* version of Positivstellensatz calculus is enough.

This is we can find polynomials g, g_i such that $p' = fg + \sum_i f_i g_i$ with degree the minimum over all derivations of p' using the inference rules.

In the proof of Theorem 1 we proceed by simplifying the refutation and then showing the lower bound for the simplified refutation. First, we can assume that all polynomials are multilinearized and we denote the multilinearized version of the polynomial p by $\overline{p}$. That is, we have

$$\overline{\prod_{i \in S} x_i^{\alpha_i}} = \prod_{i \in S} x_i \quad \text{where } \alpha_i \neq 0 \text{ for every } i \text{ in } S \tag{2}$$

This holds as the difference $p - \overline{p}$ is equal to $\sum_i f_i g_i$ for some g_i, where f_i are defined as in the beginning of the lecture, and where $\deg(f_i g_i) \leq \deg(p)$. Hence, we can make the translation without the increase in degree. We also use the following lemma which is proved in [3] as Lemma 5.2.

[3] Russell Impagliazzo, Pavel Pudlák, and Jiri Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999

Lemma 2. *For any homogeneous multilinear polynomial g with degree $d < \frac{n}{2}$, the degree of $\overline{f \cdot g}$ is exactly $d+1$, where f is the Knapsack equation.*

The lemma states that f does not cancel out when we multiply by a polynomial of small enough degree. The proof follows by considering a $\binom{n}{d+1} \times \binom{n}{d}$ matrix M indexed by subsets of $[n]$, where $M_{I,J} = 1$ if $J \subseteq I$ and 0 otherwise. It can be shown that M is a full rank matrix[4] when $d < \frac{n}{2}$. Taking a vector representing a degree d homogeneous polynomial g and multiplying it with the matrix M we get the vector representation of the polynomial $\overline{fg}$. As the matrix is full rank, the polynomial $\overline{fg}$ needs to have a monomial of degree $d+1$ with a non-zero coefficient and, hence, the degree of $\overline{fg}$ is exactly $d+1$.

[4] Daniel Henry Gottlieb. A certain class of incidence matrices. *Proceedings of the American Mathematical Society*, 17(6):1233–1237, Dec 1966

Using the previous lemma and observation, we get the following theorem (Theorem 5.1 in [5]).

[5] Russell Impagliazzo, Pavel Pudlák, and Jiri Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999

Theorem 3. *If p is derivable in degree $\lceil \frac{n}{2} \rceil$ in the equational part of* $\mathrm{PC}_>$ *then we can write p as*

$$p = fg + \sum_i f_i g_i,$$

where

- *g is multilinear, i.e., $g = \overline{g}$,*
- *$\deg(\overline{p}) = \deg(g) + 1$, and*
- *$\deg(f_i g_i) \leq \deg(p)$.*

The proof is by induction over the derivation steps. This shows that for the Knapsack problem the simpler way to deduce polynomials is as powerful as the general system, which does not hold in general. Hence, we have the following consequence of the theorem that we use in our lower bound proof.

Corollary 4. *Any refutation of the Knapsack problem of degree $d \leq \frac{n}{2}$ in Positivstellensatz calculus* $\mathrm{PC}_>$ *has the form*

$$1 + \sum_j h_j^2 = fg + \sum_i f_i g_i,$$

where the degree of h_j^2 is at most d.

The degree bound on h_j^2 follows by noting that the right hand side has degree at most d and, hence, if h_j^2 had degree greater than d there would be terms of too high degree that would not cancel out.

Now, we proceed with the usual argument where we define an operator B from monomials to real numbers and show that B is always 0 when applied to the right hand side in the Corollary 4 and for squares it is greater than or equal to 0. Hence, applying the operator to the equation in the Corollary 4 would give us $1 \leq 0$ implying that we cannot refute Knapsack in degree less than $\frac{n}{2}$ and proving Theorem 1.

Before we define the operator B, we first define numbers B_k as

$$B_k = \frac{r}{n}\frac{r-1}{n-1}\cdots\frac{r-k+1}{n-k+1}, \tag{3}$$

where in addition we define $B_0 = 1$. Then, the result of the operator B on a multilinear monomial is defined as

$$B\left(\prod_{i\in I} x_i\right) = B_{|I|}. \tag{4}$$

We extend the operator to non-multilinear monomials and polynomials in the usual way such that $B(x^2 m) = B(xm)$ and the operator is linear. The reason why we are only interested in the size of the monomial when applying the operator is that the formula is symmetric. Hence, we can change the derivation so that it has variables swapped around, which would change the content but not the degree of any term. Now, we prove some properties of the operator B.

Proposition 5. *For the operator B defined above, the following hold*

1. $B(fg) = B(f_i g_i) = 0$ *and*
2. $B(1) = 1$.

Proof. For $B(f_i g_i) = 0$ we just note that B does not take into account the exact degree of variables occurring in monomials and that $f_i = x_i^2 - x_i$. Hence,by definition B will evaluate to 0 on $(x_i^2 - x_i)g_i$. Also, $B(1) = 1$ by definition.

Now, consider $B(fX^I)$ and we have that

$$B(fX^I) = (n - |I|)B_{|I|+1} + (|I| - r)B_{|I|} = 0, \tag{5}$$

where the first equality follows by noting that for $i \in I$ the operator is equal to $B(x_i X^I) = B(X^I) = B_{|I|}$ and there are $|I|$ such indices i, and that for remaining $(n - |I|)$ indices for which $i \notin I$ we have $B(x_i X^I) = B_{|I|+1}$. Also, the free term $-r$ gets multiplied by X^I resulting in $-rB_{|I|}$ after the application of the operator B. The second equality follows by just expanding the definition of B_k. Hence, we have shown that $B(ft) = 0$ for all terms t in g and by linearity $B(fg) = 0$. □

The rest of this and the following lecture are dedicated to proving that when applying B to a square we get a non-negative value. We define the quadratic form Q on multilinear monomials X^I and X^J as $Q(X^I, X^J) = B(X^{I\cup J})$ and extend it to polynomials in the usual way. Now, we can express formally the requirement that $B(h_j^2) \geq 2$, proving which will conclude our proof.

Lemma 6. *If $\ell - 1 < r < n - \ell + 1$ then $Q(p,p) \geq 0$ for any polynomial p of degree at most ℓ.*

The proof of the lemma proceeds in the following three steps:

1. We decompose the space $P_{\leq \ell}$ of polynomials of degree at most ℓ into disjoint subspaces S_i so that we have
$$P_{\leq \ell} = S_1 \oplus S_2 \oplus S_3 \oplus \cdots \oplus S_i \oplus \cdots \tag{6}$$
2. We show that the Q, maps each subspace S_i into itself. That is, for $v \in S_i$ we have $Qv \in S_i$.
3. We show that Q is positive semidefinite in each of the subspaces S_i.

Now we proceed to decompose the space of polynomials into a chain of subspaces and define two chain operators. The factor spaces of these operators will satisfy the previous three points. We use P_t to denote the space of homogeneous polynomials of degree t for $t \leq \ell$. Now we define two types of linear operators $C_t \colon P_t \to P_{t+1}$ and $D_t \colon P_{t+1} \to P_t$. These two operators can be viewed as a sort of inverses of one another, although they are not exact inverses.

Let us first define the operator $C_t \colon P_t \to P_{t+1}$. For a given homogeneous polynomial $U = \sum_I U_I X^I$ in P_t, we define the coefficient of the monomial X^J in the polynomial $C_t(U)$ as
$$[C_t(U)]_J = \sum_{\substack{I \subset J \\ |I|=t}} U_I. \tag{7}$$
For example, if we are interested in the coefficient of $x_2x_3x_4$ in $C_t(U)$, this means that we have $J = \{2,3,4\}$ and, hence, we are summing up the coefficient of x_2x_3, x_2x_4, and x_3x_4 in the monomial U. Another example is if U is a single monomial X^I, then $C_t(X^I) = \sum_{i \notin I} x_i X^I$ as every $J \supset I$ will have the same coefficient as X^I.

The operator $D_t \colon P_{t+1} \to P_t$ is defined in the following way. Let $V = \sum_{|J|=t+1} V_J X^J$ be some polynomial in P_{t+1}. Then the coefficient of X^I in the polynomial $D_t(V)$ is defined as follows
$$[D_t(V)]_I = \sum_{\substack{J \supset I \\ |J|=t+1}} V_J. \tag{8}$$
For example, if V is a single monomial X^J, the operator D_t transforms it into $D_t(X^J) = \sum_{i \in J} X^J / x_i$ where we are taking all possible subsets I of J and assigning to X^I in $D_t(X^J)$ the same coefficient as the one that is assigned to X^J in the original polynomial V. For two concrete examples, we have that
$$D_2(x_1x_2x_4) = x_1x_2 + x_2x_4 + x_1x_4 \tag{9}$$
$$D_0(x_1 - x_2) = 1 - 1 = 0 \tag{10}$$

The subspaces we are interested in are a sequence of subspaces A_i of homogeneous polynomials of degree i defined as follows:
$$A_0 := P_0 \tag{11}$$
$$A_{t+1} := \operatorname{Ker} D_t \quad \text{for } t \geq 0. \tag{12}$$

That is, we have that a polynomial U is in A_{t+1} if and only if for every set I of size t, the sum $\sum_{J \supset I} U_J$ is equal to 0.

For a degree t polynomial $U \in A_t$ we consider polynomials of the following form

$$z_m = C_{m-1}C_{m-2}\cdots C_t(U) \quad \text{for } t < m \leq \ell, \tag{13}$$

where z_m is a homogeneous degree m polynomial. The coefficient of X^M in z_m, for $M \subseteq [n]$ and $|M| = m$, can be expressed as

$$[z_m]_M = (m-t)! \sum_{\substack{I \subset M \\ |I|=t}} U_I. \tag{14}$$

This is because there are $(m-t)$ elements that occur in M but not in I and, hence, we have $(m-t)!$ different orders of removing elements from M to reach I. Thus, we get the linear scaling factor of $(m-t)!$.

Polynomials z_m are important because in the end we want to study the polynomials we get by taking a degree t polynomial $U \in A_t$ and multiplying it by the the summation $\sum_i x_i$ raised to some power. As this summation is a part of the Knapsack constraint, finding a good representation of the resulting polynomial is helpful in finishing the proof. We can show that the polynomial we get after multilinearization is the following

$$\overline{\left(\sum_{i=1}^{n} x_i\right)^{m-t} U} = C_{m-1}C_{m-2}\cdots C_t(U) + \\ + \alpha_{m-1}C_{m-2}C_{m-3}\cdots C_t(U) + \cdots + \\ + \alpha_i C_i \cdots C_t(U) + \cdots + \alpha_{t+1}U, \tag{15}$$

where α_i are coefficients that depend only on m and t and not on the concrete polynomial U. The proof is just by a simple expansion of the resulting polynomial where we do not need to care about the concrete values of α_i. Hence, it follows that U can be expanded as the sum of different C_i operators applied to U.

References

[Got66] Daniel Henry Gottlieb. A certain class of incidence matrices. *Proceedings of the American Mathematical Society*, 17(6):1233–1237, Dec 1966.

[Gri01] Dima Grigoriev. Complexity of positivstellensatz proofs for the knapsack. *computational complexity*, 10(2):139–154, 2001.

[IPS99] Russell Impagliazzo, Pavel Pudlák, and Jiri Sgall. Lower bounds for the polynomial calculus and the Gröbner basis algorithm. *Computational Complexity*, 8(2):127–144, 1999.