Gröbner problems from tensor rank

Lars Hellström

Version 0.2, 1 March 2006

Abstract

The following Gröbner basis problems trivially arise from the question of whether a particular tensor has a decomposition into at most some number T of simple tensors. Of primary interest is to decide whether the ideal is trivial, but secondarily one also wants to find some point in the corresponding variety.

Two typical features of these problems which make them difficult is that they have quite a lot of variables and large symmetry groups. On the other hand, the degrees are about as low as they can get without the system becoming linear.

1 The examples

For the hurried reader, here are the benchmark examples. I will in general write $[n]$ for the set $\{1,\dots,n\}$, with in particular $[1]=\{1\}$ and $[0]=\varnothing$.

Example 1 (identity matrix rank). This family of examples have two parameters n and T. There are $2nT$ variables called x_{ik} and y_{ik} respectively for $i \in [n]$ and $k \in [T]$. There are n^2 equations, namely

$$\sum_{k=1}^{T} x_{ik} y_{jk} = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i \neq j, \end{cases} \qquad \text{for all } i,j \in [n]. \tag{1}$$

Regardless of the base field, this system has no solutions for $T<n$ and plenty of solutions whenever $T \geqslant n$.

Example 1 was constructed mainly to be the classical "minimal example": a simplified problem still exhibiting all the nastiness of the problem one really wants to solve. My real problem is that of Example 2.

Example 2 (matrix multiplication tensor rank). This family of examples have four parameters l, m, n, and T, but the roles of the first three are interchangeable. There are $(lm+ln+mn)T$ variables x_{ijk}, y_{jrk}, and z_{irk} for $i \in [l]$, $j \in [m]$, $r \in [n]$, and $k \in [T]$. The system of equations they have to satisfy is

$$\sum_{k=1}^{T} x_{ijk} y_{rsk} z_{uvk} = \begin{cases} 1 & \text{if } i=u,\ j=r, \text{ and } s=v, \\ 0 & \text{otherwise,} \end{cases}$$
$$\text{for all } i,u \in [l],\ j,r \in [m], \text{ and } s,v \in [n]; \tag{2}$$

in total $(lmn)^2$ different equations.

As in the previous example there is for all $l, m, n \in \mathbb{Z}_{>0}$ a threshold value $R(l, m, n)$ for T such that solutions exist for all $T \geqslant R(l, m, n)$ (solutions for larger T can be manufactured by padding with zeroes) but not for any $T < R(l, m, n)$. Some known facts are:

1. $R(l, m, n) \leqslant lmn$.
2. $R(2, 2, 2) = 7$ [Str69, HK71].
3. $R(2, 2, 3) = 11$ [HK71].
4. $R(2, 3, 3) = 15$ [HK71].
5. $19 \leqslant R(3, 3, 3) \leqslant 23$ [Blä03, Lad76].

It is believed by some that R depends on the characteristic of the underlying field, but as far as I can tell noone has produced an example of this. On the other hand, R is exactly known only for very few values of (l, m, n).

2 Matrix rank

Now what is actually going on here? Example 1 is all about decomposing the $n \times n$ identity matrix as a sum

$$\mathbf{x}_1 \mathbf{y}_1^{\mathrm{T}} + \cdots + \mathbf{x}_T \mathbf{y}_T^{\mathrm{T}} \tag{3}$$

where each $\mathbf{x}_k$ and $\mathbf{y}_k$ is a (column) n-vector (and the transpose thus turns $\mathbf{y}_k$ into row vectors, so that each term is an $n \times 1$ by $1 \times n$ matrix product). This succeeds if and only if $T \geqslant n$, because the property of a matrix A to possess a decomposition of the form (3) is equivalent to the claim that the rank of A is $\leqslant T$; the threshold for having such a decomposition is in fact the coordinate-free definition of matrix rank.

[Fill in: Describe the **symmetries** of the system.]

3 3-tensor rank

For a vector space U, denote by U^* the linear dual of U.

A matrix or linear transformation $U \longrightarrow V$ can be viewed as a 2-tensor, i.e., an element in a tensor product $U^* \otimes V$ of two (finite dimensional) vector spaces. In the same way, a *bilinear map* $U \times V \longrightarrow W$ can be viewed as an element of the tensor triple product $U^* \otimes V^* \otimes W$, or for short as a 3-tensor. As an example of this, consider complex multiplication as an operation over $\mathbb{R}$. The vector space $\mathbb{C}$ has the standard basis $\{1, i\}$, whereas its dual $\mathbb{C}^*$ has the standard basis $\{\mathrm{Re}, \mathrm{Im}\}$. From the definition $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$, one can read off that the complex multiplication tensor is

$$\mathrm{Re} \otimes \mathrm{Re} \otimes 1 - \mathrm{Im} \otimes \mathrm{Im} \otimes 1 + \mathrm{Re} \otimes \mathrm{Im} \otimes i + \mathrm{Im} \otimes \mathrm{Re} \otimes i \tag{4}$$

which in a more array-like manner can be written as for example

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

Since this three index tensor has two full rank matrices as components no matter how the indices are arranged, one may think that the decomposition (4) uses the minimal number of terms, but in fact it does not. There is a three term decomposition

$$(\mathrm{Re} + \mathrm{Im}) \otimes (\mathrm{Re} + \mathrm{Im}) \otimes i + \mathrm{Re} \otimes \mathrm{Re} \otimes (1 - i) - \mathrm{Im} \otimes \mathrm{Im} \otimes (1 + i) \qquad (5)$$

of this tensor which demonstrates that the rank is at most 3, and in fact the rank is exactly three.

[Write lots of more things.]

References

[Blä03] Markus Bläser. On the complexity of the multiplication of matrices of small formats. *J. Complexity*, 19(1):43–60, 2003.

[HK71] J. E. Hopcroft and L. R. Kerr. On minimizing the number of multiplications necessary for matrix multiplication. *SIAM J. Appl. Math.*, 20:30–36, 1971.

[Lad76] Julian D. Laderman. A noncommutative algorithm for multiplying 3×3 matrices using 23 muliplications. *Bull. Amer. Math. Soc.*, 82(1):126–128, 1976.

[Str69] Volker Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13:354–356, 1969.