



HAL
open science

On the reduction of multivariate quadratic systems to best rank-1 approximation of three-way tensors

Alex P da Silva, Pierre Comon, André L F de Almeida

► **To cite this version:**

Alex P da Silva, Pierre Comon, André L F de Almeida. On the reduction of multivariate quadratic systems to best rank-1 approximation of three-way tensors. *Applied Mathematics Letters*, 2016, 62, pp.9 - 15. 10.1016/j.aml.2016.06.006 . hal-01567408

HAL Id: hal-01567408

<https://hal.science/hal-01567408>

Submitted on 23 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the reduction of multivariate quadratic systems to best rank-1 approximation of three-way tensors

Alex P. da Silva¹, Pierre Comon¹, André L. F. de Almeida²

Abstract

In this paper, we show that a general quadratic multivariate system in the real field can be reduced to a best rank-1 three-way tensor approximation problem. This fact provides a new approach to tackle a system of quadratic polynomial equations. Some experiments using the standard alternating least squares (ALS) algorithm are drawn to evince the usefulness of rank-1 tensor approximation methods.

Keywords: quadratic systems, tensor, rank-1 approximation.

1. Introduction

Let P1 be the general system of quadratic polynomial equations given by

$$\mathbf{P1:} \begin{cases} \mathbf{x}^\top \mathbf{A}_1 \mathbf{x} + \mathbf{b}_1^\top \mathbf{x} + c_1 = 0 \\ \vdots \\ \mathbf{x}^\top \mathbf{A}_m \mathbf{x} + \mathbf{b}_m^\top \mathbf{x} + c_m = 0, \end{cases}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{A}_j \in \mathbb{R}^{n \times n}$ are symmetric matrices, $\mathbf{b}_j \in \mathbb{R}^n$ and $c_j \in \mathbb{R}$, $1 \leq j \leq m$.

To tackle this problem, some mathematical techniques can be applied, such as Newton and tensor-based algorithms [1, 2], Gröbner bases, resultants and eigenvalues/eigenvectors of companion matrices [3], semidefinite relaxations [4, 5, 6], numerical homotopy [7, 8], low-rank matrix recovery [9], and symbolic computation [10]. The complexity class of System (P1) is NP-hard.

This problem is of great interest in various applications. For instance, in game theory, the Nash equilibria of a non-cooperative game between two players can be found by solving a multivariate quadratic system [11]. In cryptography, the security of systems depends on the difficulty to solve large quadratic systems in finite fields [12, 13]. In [14] the authors present the design of multivariate filter banks modeled by quadratic polynomial systems. A last example of application lies in multilinear algebra [15], where the authors propose an efficient algorithm to decompose a symmetric tensor and show the equivalence between an existence condition of the decomposition and the solution of a system of quadratic equations.

Let $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a three-way tensor with entries in the real field. The best rank-1 approximation of \mathcal{T} can be obtained by solving the following optimization problem

$$p^* = \min_{\mathbf{a}_i, 1 \leq i \leq 3} \|\mathcal{T} - \mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \mathbf{a}_3\|, \quad (1)$$

where $\mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \mathbf{a}_3$ is a rank-1 tensor with approximating factors $\mathbf{a}_i \in \mathbb{R}^{n_i}$, $1 \leq i \leq 3$, \otimes is the tensor product, and $\|\cdot\|$ is the Frobenius norm. The solution of Problem (1) always exists in \mathbb{R} , since the set of rank-1 tensors is closed [16]. It is

¹GIPSA-Lab, CNRS UMR2016, Grenoble Campus, BP.46, F-38402 Saint Martin d'Herès CEDEX, France (alex.pereira-da-silva@gipsa-lab.grenoble-inp.fr, pierre.comon@gipsa-lab.grenoble-inp.fr)

²Department of Teleinformatics Engineering, Federal University of Ceará, Campus do Pici, s/n, CP 6005 60455-970, Fortaleza, CE, Brazil (andre@gtel.ufc.br)

known that standard tensor approximation algorithms, such as alternating least squares (ALS) [17] generally deliver satisfactory solutions in practice in reasonable time for the best rank-1 approximation problem. Other methods to solve this problem can be found in [6, 18].

The goal of this paper is to show that multivariate quadratic systems can be reduced to best rank-1 three-way tensor approximation problems, which provides a new approach with a large amount of tensor tools to deal with System (P1).

2. From Quadratic Systems To Best Rank-1 Approximations

Let \mathcal{J} be the ideal in $\mathbb{R}[\mathbf{x}]$ generated by the set of polynomials of (P1), that is, $\mathcal{J} = \langle h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_m(\mathbf{x}) \rangle$, where $h_j(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}_j \mathbf{x} + \mathbf{b}_j^\top \mathbf{x} + c_j$, $1 \leq j \leq m$. Thus, the set of solutions of (P1) is defined by the affine variety $V(\mathcal{J}) := \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = 0, \forall f \in \mathcal{J}\}$. Also define the optimization problem (P2) as follows

$$\mathbf{P2:} \quad p_2^* = \min_{\|\mathbf{y}\|=1} p(\mathbf{y})$$

where $p(\mathbf{y}) = \sum_{j=1}^m (\mathbf{y}^\top \mathbf{Q}_j \mathbf{y})^2$, for $\mathbf{y} \in \mathbb{R}^{n+1}$ and

$$\mathbf{Q}_j = \begin{bmatrix} \mathbf{A}_j & \mathbf{b}_j/2 \\ \mathbf{b}_j^\top/2 & c_j \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}.$$

Define the set of global minimizers of (P2) as $\mathcal{S}_{P2} = \{\mathbf{y} \in \mathbb{R}^{n+1} \mid p(\mathbf{y}) = p_2^*, \|\mathbf{y}\| = 1\}$, and the subset $\bar{\mathcal{S}}_{P2} \subseteq \mathcal{S}_{P2}$ given by $\bar{\mathcal{S}}_{P2} = \mathcal{S}_{P2} \cap (\mathbb{R}^n \times \mathbb{R} \setminus \{0\})$. In other words, $\bar{\mathcal{S}}_{P2}$ is the set of solutions of Problem (P2) such that $y_{n+1} \neq 0$. Let also $\mathcal{N} = \{\mathbf{z} \in \mathbb{R}^{n+1} \mid \mathbf{z} = \mathbf{y}/y_{n+1}, \forall \mathbf{y} \in \bar{\mathcal{S}}_{P2}\}$. Proposition below connects Problems (P1) and (P2).

Proposition 1. *If $V(\mathcal{J}) \neq \emptyset$ then $V(\mathcal{J}) \times \{1\} = \mathcal{N}$.*

Proof. By setting $\mathbf{y} = [\mathbf{x} \ 1]^\top$, it turns out that

$$\mathbf{x}^\top \mathbf{A}_j \mathbf{x} + \mathbf{b}_j^\top \mathbf{x} + c_j = \mathbf{y}^\top \mathbf{Q}_j \mathbf{y},$$

$\forall j \in \{1, 2, \dots, m\}$. This shows that the set of solutions of the following system, equivalent to (P1),

$$\begin{cases} \mathbf{y}^\top \mathbf{Q}_j \mathbf{y} = 0, \text{ for } 1 \leq j \leq m \\ y_{n+1} = 1 \end{cases} \quad (2)$$

is $V(\mathcal{J}) \times \{1\}$. Now, consider the optimization problem (P2). Since $p(\mathbf{y})$ is by construction the sum of squares of the quadratic polynomials $\mathbf{y}^\top \mathbf{Q}_j \mathbf{y}$, $1 \leq j \leq m$, and $V(\mathcal{J}) \neq \emptyset$, it follows that $p_2^* = 0$. Thus, $\forall \bar{\mathbf{y}} \in V(\mathcal{J}) \times \{1\}$, $p(\bar{\mathbf{y}}/\|\bar{\mathbf{y}}\|) = 0$, which implies that $\bar{\mathbf{y}}/\|\bar{\mathbf{y}}\| \in \bar{\mathcal{S}}_{P2}$ and $\bar{\mathbf{y}} \in \mathcal{N}$. This proves that $V(\mathcal{J}) \times \{1\} \subseteq \mathcal{N}$.

On the other hand, $\forall \mathbf{y} \in \bar{\mathcal{S}}_{P2}$, $p(\mathbf{y}/y_{n+1}) = 0$, which implies that $(\mathbf{y}/y_{n+1})^\top \mathbf{Q}_j (\mathbf{y}/y_{n+1}) = 0$, $1 \leq j \leq m$. Thus, \mathbf{y}/y_{n+1} is solution of (2), which leads to $\mathcal{N} \subseteq V(\mathcal{J}) \times \{1\}$, and the proof is complete. \square

Proposition 1 says that all solutions of System (P1) can be recovered from Problem (P2) by taking the n first elements of \mathbf{y}/y_{n+1} , $\forall \mathbf{y} \in \bar{\mathcal{S}}_{P2}$.

Remark 1. *When $\mathbf{b}_j = \mathbf{0}$, $c_j = 0$, $\forall j \in \{1, 2, \dots, m\}$, $V(\mathcal{J}) - \{\mathbf{0}\}$ can be viewed as a projective variety, since $\mathbf{x} \in V(\mathcal{J}) \implies \alpha \mathbf{x} \in V(\mathcal{J}) \forall \alpha \in \mathbb{R}$. Therefore, it is enough to solve the system for $\|\mathbf{x}\| = 1$. By setting $\mathbf{y} = \mathbf{x}$, and $\mathbf{Q}_j = \mathbf{A}_j$, $1 \leq j \leq m$, all nontrivial solutions of (P1) can be recovered by solving (P2).*

Now, consider the following optimization problem

$$\mathbf{P3:} \quad p_3^* = \max_{\|\mathbf{y}\|=1} (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M} (\mathbf{y} \boxtimes \mathbf{y})$$

where $\mathbf{M} \in \mathbb{R}^{(n+1)^2 \times (n+1)^2}$ is a positive semidefinite matrix given by $\mathbf{M} = \lambda_{\max} \mathbf{I} - \mathbf{M}_Q$, where $\mathbf{M}_Q = \sum_{j=1}^m \mathbf{Q}_j \boxtimes \mathbf{Q}_j$, \mathbf{I} is the identity matrix, λ_{\max} is the largest eigenvalue of \mathbf{M}_Q , and \boxtimes denotes the Kronecker product.

Proposition 2. (P2) \iff (P3).

Proof. Since

$$p(\mathbf{y}) = \sum_{j=1}^m (\mathbf{y}^\top \mathbf{Q}_j \mathbf{y})^2 = \sum_{j=1}^m (\mathbf{y} \boxtimes \mathbf{y})^\top (\mathbf{Q}_j \boxtimes \mathbf{Q}_j) (\mathbf{y} \boxtimes \mathbf{y}) = (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M}_Q (\mathbf{y} \boxtimes \mathbf{y}),$$

it follows that

$$\begin{aligned} \min_{\|\mathbf{y}\|=1} (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M}_Q (\mathbf{y} \boxtimes \mathbf{y}) &= \max_{\|\mathbf{y}\|=1} -(\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M}_Q (\mathbf{y} \boxtimes \mathbf{y}) \iff \max_{\|\mathbf{y}\|=1} \lambda_{\max} - (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M}_Q (\mathbf{y} \boxtimes \mathbf{y}) \\ &= \max_{\|\mathbf{y}\|=1} (\mathbf{y} \boxtimes \mathbf{y})^\top (\lambda_{\max} \mathbf{I} - \mathbf{M}_Q) (\mathbf{y} \boxtimes \mathbf{y}) = \max_{\|\mathbf{y}\|=1} (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M} (\mathbf{y} \boxtimes \mathbf{y}). \quad \square \end{aligned}$$

Remark 2. Problem (P3) can be viewed as a symmetric best rank-1 approximation of a fourth order tensor. Indeed, it can be written as $\min_{\|\mathbf{y}\|=1} \|\mathbf{M} - \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{y}\|^2$, where $\mathbf{M} \in \mathbb{R}^{n+1 \times n+1 \times n+1 \times n+1}$ is a symmetric tensor constructed from \mathbf{M} . This problem is difficult to solve and some standard algorithms, such as ALS, could be adapted to deal with it. However, ignoring the symmetry generally leads to faster convergence. For this reason, reducing the system to a standard three-way rank-1 approximation problem is attractive under the algorithmic point of view.

In the following, we establish a link between (P3) and a rank-1 three-way tensor approximation problem. Let $(\lambda_k, \mathbf{v}_k)$, $1 \leq k \leq K$, $K = \text{rank}\{\mathbf{M}\}$, be the k -th eigenpair of \mathbf{M} . Due to the semi-definiteness of \mathbf{M} , we can define real vectors $\mathbf{t}_k = \sqrt{\lambda_k} \mathbf{v}_k \in \mathbb{R}^{(n+1)^2}$. Let Unvec be the operator that *unstacks* $n+1$ by $n+1$ the elements of an $(n+1)^2$ vector, and reshapes them as columns of an $n+1 \times n+1$ matrix.

Thus, we set $\mathbf{T}_k = \text{Unvec}(\mathbf{t}_k)$, $1 \leq k \leq \text{rank}\{\mathbf{M}\}$. The matrices \mathbf{T}_k can be viewed as frontal slices of a three-way tensor $\mathcal{T} \in \mathbb{R}^{n+1 \times n+1 \times K}$ [19], as illustrated in the figure below.

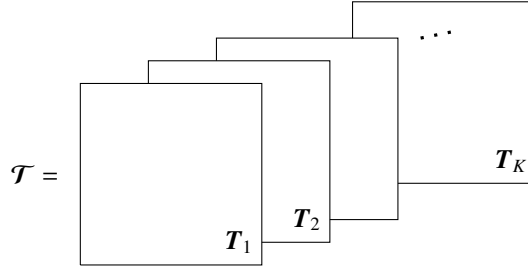


Figure 1: Frontal slices of a three-way tensor $\mathcal{T} \in \mathbb{R}^{n+1 \times n+1 \times K}$.

Define the following rank-1 tensor approximation problem with two identical factors.

$$\mathbf{P4:} \quad p_4^* = \min_{\|\mathbf{y}\|=1, \mathbf{w}} \|\mathcal{T} - \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{w}\|.$$

The following proposition holds.

Proposition 3. (P3) \iff (P4).

Proof. Expand matrix \mathbf{M} into its eigencomponents as: $\mathbf{M} = \sum_{k=1}^K \lambda_k \mathbf{v}_k \mathbf{v}_k^\top = \sum_{k=1}^K \mathbf{t}_k \mathbf{t}_k^\top$, and write the objective function of (P3) as $(\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{M} (\mathbf{y} \boxtimes \mathbf{y}) = \sum_{k=1}^K (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{t}_k \mathbf{t}_k^\top (\mathbf{y} \boxtimes \mathbf{y})$. Set $w_k = (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{t}_k$. Problem (P3) can be rewritten as

$$\max_{(\mathbf{y}, \mathbf{w}) \in \mathcal{C}} \sum_{k=1}^K w_k \mathbf{t}_k^\top (\mathbf{y} \boxtimes \mathbf{y}) \quad (3)$$

where $\mathcal{C} = \{(\mathbf{y}, \mathbf{w}) \in \mathbb{R}^n \times \mathbb{R}^K, : \|\mathbf{y}\| = 1, \mathbf{w} = [w_1 \ w_2 \ \dots \ w_K]^\top, w_k = (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{t}_k, 1 \leq k \leq K\}$. Next, we have the equivalence

$$\max_{(\mathbf{y}, \mathbf{w}) \in \mathcal{C}} \sum_{k=1}^K w_k \mathbf{t}_k^\top (\mathbf{y} \boxtimes \mathbf{y}) \iff \min_{(\mathbf{y}, \mathbf{w}) \in \mathcal{C}} \sum_{k=1}^K \mathbf{t}_k^\top \mathbf{t}_k - w_k \mathbf{t}_k^\top (\mathbf{y} \boxtimes \mathbf{y}).$$

Yet, since $(\mathbf{y}, \mathbf{w}) \in \mathcal{C}$, it turns out that $\mathbf{t}_k^\top \mathbf{t}_k - w_k \mathbf{t}_k^\top (\mathbf{y} \boxtimes \mathbf{y}) = \|\mathbf{t}_k - w_k (\mathbf{y} \boxtimes \mathbf{y})\|^2$, $1 \leq k \leq K$. Therefore, Problem (P3) is equivalent to

$$\min_{(\mathbf{y}, \mathbf{w}) \in \mathcal{C}} \sum_{k=1}^K \|\mathbf{t}_k - w_k (\mathbf{y} \boxtimes \mathbf{y})\|^2 \quad (4)$$

or, to minimize the Lagrangian:

$$\Upsilon(\mathbf{y}, \mathbf{w}) = \sum_{k=1}^K \|\mathbf{t}_k - w_k (\mathbf{y} \boxtimes \mathbf{y})\|^2 + \eta (\|\mathbf{y}\| - 1) + \sum_{k=1}^K \gamma_k (w_k - (\mathbf{y} \boxtimes \mathbf{y})^\top \mathbf{t}_k),$$

where γ_k , $1 \leq k \leq K$, and η are Lagrangian multipliers.

We now need to relax the constraint on w_k . Stationary points w.r.t. w_k yield

$$\frac{\partial \Upsilon(\mathbf{y}, \mathbf{w})}{\partial w_k} = -2(\mathbf{t}_k - w_k (\mathbf{y} \boxtimes \mathbf{y}))^\top (\mathbf{y} \boxtimes \mathbf{y}) + \gamma_k = 0 \implies (-\mathbf{t}_k + w_k (\mathbf{y} \boxtimes \mathbf{y}))^\top (\mathbf{y} \boxtimes \mathbf{y}) + \gamma_k/2 = 0.$$

Now using all the constraints defined in \mathcal{C} (*i.e.* stationary equations w.r.t. γ_k and η), the latter equation leads to $\gamma_k = 0$, $\forall k$, which means that the constraints on w_k do not need to be imposed. Hence we end up with the simplified Lagrangian:

$$\Upsilon(\mathbf{y}, \mathbf{w}) = \sum_{k=1}^K \|\mathbf{t}_k - w_k (\mathbf{y} \boxtimes \mathbf{y})\|^2 + \eta (\|\mathbf{y}\| - 1). \quad (5)$$

To complete the proof, note that by reshaping the entries of the vectors \mathbf{t}_k using the Unvec operator, (5) can still be written with the help of matrix slices:

$$\min_{\|\mathbf{y}\|=1, \mathbf{w}} \sum_{k=1}^K \|\mathbf{T}_k - w_k \mathbf{y} \mathbf{y}^\top\|^2 \quad (6)$$

or in its tensor representation

$$\min_{\|\mathbf{y}\|=1, \mathbf{w}} \|\mathcal{T} - \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{w}\|^2. \quad \square$$

Problem (P4) is a best partially symmetric rank-1 approximation, which cannot be directly tackled by standard tensor approximation algorithms. However, this problem can still be changed in order to drop the constraint of identical factors.

According to [20], the best rank-1 approximation of a real symmetric tensor with respect to a subset of indices (modes) can be chosen symmetric with respect to this subset (see Theorem 1 therein). Thus, if tensor \mathcal{T} in Problem (P4) were symmetric with respect to the first two modes, we would not need to impose any constraint on the factors, so that the best rank-1 approximation could be chosen partially symmetric.

In general, \mathcal{T} is not symmetric with respect to the first two modes. However, the linear transformation $\mathbf{T}'_k = (\mathbf{T}_k + \mathbf{T}_k^\top)/2$ on each k -th slice \mathbf{T}_k of \mathcal{T} , ensures that any best partially symmetric rank-1 approximation of \mathcal{T} is actually a best rank-1 approximation of a tensor \mathcal{T}' with slices \mathbf{T}'_k (which is therefore symmetric with respect to the first two modes). In other words, Problem (P4) can be reduced to the problem

$$\mathbf{P5:} \quad p_5^* = \min_{\|\mathbf{y}\|=1, \mathbf{x}, \mathbf{w}} \|\mathcal{T}' - \mathbf{x} \otimes \mathbf{y} \otimes \mathbf{w}\|.$$

Before proving this, we need the following lemma.

Lemma 4. *Let $f(\mathbf{X}, \alpha) = \sum_{k=1}^K \|\mathbf{A}_k - \alpha_k \mathbf{X}\|^2$ and $g(\mathbf{X}, \alpha) = \sum_{k=1}^K \|(\mathbf{A}_k + \mathbf{A}_k^\top)/2 - \alpha_k \mathbf{X}\|^2$ with $\mathbf{X}, \mathbf{A}_k \in \mathbb{R}^{n \times n}$, $1 \leq k \leq K$, $\alpha = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_K]^\top$. If \mathbf{X} is symmetric, then*

$$\arg \min_{\mathbf{X}, \alpha} f(\mathbf{X}, \alpha) = \arg \min_{\mathbf{X}, \alpha} g(\mathbf{X}, \alpha).$$

Proof. Let $h(\mathbf{X}, \boldsymbol{\alpha}) = \sum_{k=1}^K \|\mathbf{A}_k^\top - \alpha_k \mathbf{X}\|^2$. Since \mathbf{X} is symmetric, it follows that

$$\arg \min_{\mathbf{X}, \boldsymbol{\alpha}} f(\mathbf{X}, \boldsymbol{\alpha}) = \arg \min_{\mathbf{X}, \boldsymbol{\alpha}} h(\mathbf{X}, \boldsymbol{\alpha}) = \arg \min_{\mathbf{X}, \boldsymbol{\alpha}} f(\mathbf{X}, \boldsymbol{\alpha}) + h(\mathbf{X}, \boldsymbol{\alpha}).$$

By expanding the functions,

$$f(\mathbf{X}, \boldsymbol{\alpha}) + h(\mathbf{X}, \boldsymbol{\alpha}) = 2\|\mathbf{X}\|^2 \sum_{k=1}^K \alpha_k^2 - 2 \sum_{k=1}^K \alpha_k \text{tr}\{(\mathbf{A}_k + \mathbf{A}_k^\top)\mathbf{X}\} + 2 \sum_{k=1}^K \|\mathbf{A}_k\|^2, \text{ and}$$

$$g(\mathbf{X}, \boldsymbol{\alpha}) = \|\mathbf{X}\|^2 \sum_{k=1}^K \alpha_k^2 - \sum_{k=1}^K \alpha_k \text{tr}\{(\mathbf{A}_k + \mathbf{A}_k^\top)\mathbf{X}\} + \frac{1}{4} \sum_{k=1}^K \|\mathbf{A}_k + \mathbf{A}_k^\top\|^2,$$

where tr is the trace operator. Thus, $f(\mathbf{X}, \boldsymbol{\alpha}) + h(\mathbf{X}, \boldsymbol{\alpha}) = 2g(\mathbf{X}, \boldsymbol{\alpha}) + c$, for some constant c , and the proof is complete. \square

Proposition 5. Let \mathcal{S}_{P4} and \mathcal{S}_{P5} be the sets of rank-1 tensors that minimize Problems (P4) and (P5), respectively. Then $\mathcal{S}_{P4} \subseteq \mathcal{S}_{P5}$.

Proof. Since $\mathbf{y}\mathbf{y}^\top$ is symmetric, we can apply Lemma 4 to the cost function of the slice form of Problem (P4) to show that

$$\arg \min_{\mathbf{y}, \mathbf{w}} \sum_{k=1}^K \|\mathbf{T}_k - w_k \mathbf{y}\mathbf{y}^\top\|^2 = \arg \min_{\mathbf{y}, \mathbf{w}} \sum_{k=1}^K \|\mathbf{T}'_k - w_k \mathbf{y}\mathbf{y}^\top\|^2.$$

That means that any rank-1 tensor $\mathcal{X} = \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{w} \in \mathcal{S}_{P4}$ is a partially symmetric rank-1 tensor that minimizes (P5), which is in turn a best rank-1 approximation because of Theorem 1 in [20]. \square

3. Discussion

In this section, we present some numerical experiments at which a solution of generic quadratic systems are computed by using the standard ALS algorithm for rank-1 tensor approximations. A total of 500 systems for each scenario (n, m) was generated with the entries of $\mathbf{A}_j, \mathbf{b}_j, c_j, 1 \leq j \leq m$, distributed according to a uniform measure in $[-1, 1]$. To initialize ALS algorithm, we have employed a partially symmetric random initialization, which consists of the minimizer of the approximation error $\|\mathcal{T}' - \mathcal{X}\|$ among a sample of 10000 partially symmetric tensors $\mathcal{X} = \mathbf{y} \otimes \mathbf{y} \otimes \mathbf{w}$, with the entries of vectors \mathbf{y} and \mathbf{w} also uniformly distributed in $[-1, 1]$. Table 1 summarizes the percentage p of systems in which a solution was successfully obtained, and the average computational time for initialization, construction of \mathcal{T}' , and the rank-1 approximation itself. We assume that a solution satisfies a system if $\max(|\mathbf{y}^\top \mathbf{Q}_1 \mathbf{y}|, |\mathbf{y}^\top \mathbf{Q}_2 \mathbf{y}|, \dots, |\mathbf{y}^\top \mathbf{Q}_m \mathbf{y}|) \leq 0.001$. All time measurements are displayed in seconds. The experiments were performed with the software Matlab[®] in a MAC OSX 10.8.5 with processor 3.2 GHz Intel core i5, and memory 8G 1600MHz DDR3.

Table 1: Performance of the tensor approximation ALS algorithm to extract one solution of generic quadratic systems.

Scenario (n, m)	Initialization Time	Time for constructing \mathcal{T}'	Time ALS	Total Time	p (%)
(2, 3)	1.2707	8.5767e-04	0.0852	1.3567	100
(4, 4)	1.5190	0.0014	2.0137	3.5340	100
(5, 4)	1.5288	0.0024	0.4613	1.9925	100
(6, 6)	1.5800	0.0035	3.5453	5.1288	100
(7, 9)	1.6444	0.0051	0.9939	2.6433	98.6
(10, 8)	2.7514	0.0133	2.0391	4.8037	100
(10, 10)	2.7705	0.0133	8.8520	11.6358	100
(20, 20)	11.8718	0.1993	51.4871	63.5582	100

Some conclusions and remarks about the rank-1 approximation approach:

- Although a solution of (P5) is not necessarily a solution of (P1), the experiments showed that ALS algorithm delivers a solution of (P1) in general. Indeed, only in the over-determined scenario (7, 9) a solution was not found for a few systems, but the reason was that ALS did not converge to a best rank-1 approximation;
- Routine simulations have shown that ALS delivers partially symmetric rank-1 approximations for partially symmetric tensors;
- The ALS algorithm is easy to implement and converges fast for generic tensors approximations. Additionally, ALS has also local linear convergence properties. If the starting point is in the neighborhood of a global solution of (P5) then the iterates of ALS converge linearly to this solution [21];
- The results also confirmed that a solution of a quadratic polynomial system can be extracted in a few seconds with the rank-1 tensor approach. However, the computational time becomes large for scenarios $(n, m) \geq (10, 10)$. For instance, when $n = 10$, the tensor \mathcal{T}' can have dimensions up to $11 \times 11 \times 121$.
- From Table 1, the computational time of the ALS algorithm is larger for systems having as many variables as equations. For instance, compare scenarios (4, 4) and (6, 6) with scenarios (5, 4) and (7, 9), respectively;
- Despite the fact that the construction of \mathcal{T}' depends on the number of equations m , Problem (P5) itself does not. This is an advantage compared to other methods that take into account the number of equations.

Acknowledgments

This work has been funded by the European Research Council under the 7th Framework Program FP7/2007-2013 Grant Agreement no. 320594, and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under the program Ciências sem Fronteiras-Brazil.

References

- [1] J. E. Dennis Jr, R. B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations, Vol. 16, Siam, 1996.
- [2] R. B. Schnabel, P. D. Frank, Tensor methods for nonlinear equations, SIAM Journal on Numerical Analysis 21 (5) (1984) 815–843.
- [3] D. A. Cox, J. Little, D. O’Shea, Using algebraic geometry, Vol. 185, Springer Science & Business Media, 2006.
- [4] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, Mathematical programming 96 (2) (2003) 293–320.
- [5] J. B. Lasserre, Global optimization with polynomials and the problem of moments, SIAM Journal on Optimization 11 (3) (2001) 796–817.
- [6] M. A. Bucero, B. Mourrain, Border basis relaxation for polynomial optimization, Journal of Symbolic Computation 74 (2016) 378–399.
- [7] T.-Y. Li, Numerical solution of multivariate polynomial systems by homotopy continuation methods, Acta numerica 6 (1997) 399–436.
- [8] J. Verschelde, Polynomial homotopies for dense, sparse and determinantal systems, Mathematical Systems Research Institute (MSRI), preprint # 1999-041.
- [9] M. Davenport, J. Romberg, An overview of low-rank matrix recovery from incomplete observations, IEEE Journal of Selected Topics in Signal Processing 10 (2016) 608 – 622.
- [10] D. Grigoriev, D. V. Pasechnik, Polynomial-time computing over quadratic maps i: sampling in real algebraic sets, computational complexity 14 (1) (2005) 20–52.
- [11] R. J. Lipton, E. Markakis, Nash equilibria via polynomial equations, in: LATIN 2004: Theoretical Informatics, Springer, 2004, pp. 413–422.
- [12] N. Courtois, A. Klimov, J. Patarin, A. Shamir, Efficient algorithms for solving overdefined systems of multivariate polynomial equations, in: Advances in Cryptology, EUROCRYPT 2000, Springer, 2000, pp. 392–407.
- [13] E. Thomae, C. Wolf, Solving underdetermined systems of multivariate quadratic equations revisited, in: Public Key Cryptography–PKC 2012, Springer, 2012, pp. 156–171.
- [14] Z. Lin, L. Xu, Q. Wu, Applications of gröbner bases to signal and image processing: A survey, Linear algebra and its applications 391 (2004) 169–202.
- [15] J. Brachat, P. Comon, B. Mourrain, E. Tsigaridas, Symmetric tensor decomposition, Linear Algebra and its Applications 433 (11) (2010) 1851–1872.
- [16] V. De Silva, L.-H. Lim, Tensor rank and the ill-posedness of the best low-rank approximation problem, SIAM Journal on Matrix Analysis and Applications 30 (3) (2008) 1084–1127.
- [17] N. Li, S. Kindermann, C. Navasca, Some convergence results on the regularized alternating least-squares method for tensor decomposition, Linear Algebra and its Applications 438 (2) (2013) 796–812.
- [18] J. Nie, L. Wang, Semidefinite relaxations for best rank-1 tensor approximations, SIAM Journal on Matrix Analysis and Applications 35 (3) (2014) 1155–1179.
- [19] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM review 51 (3) (2009) 455–500.
- [20] S. Friedland, Best rank one approximation of real symmetric tensors can be chosen symmetric, Frontiers of Mathematics in China 8 (1) (2013) 19–40.
- [21] A. Uschmajew, Local convergence of the alternating least squares algorithm for canonical tensor approximation, SIAM Jour. matrix Analysis 33 (2) (2012) 639–652.