with the Department of Engineering, King's College London, London, WC2R 2LS, U.K. Her research interests include information theory, statistical learning theory, stochastic control, and optimization.

Osvaldo Simeone (osvaldo.simeone@kcl.ac.uk) is a professor of information engineering with the Centre for Telecommunications Research in the Department of Informatics, King's College London, London, WC2R 2LS, U.K., where he directs the King's Communications, Learning, and Information Processing lab. He is a Fellow of IEEE.

References

[1] S. Shalev-Shwartz, "Online learning and online convex optimization," *Found. Trends Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2012. doi: 10.1561/2200000018.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey: Wiley-Interscience, July 2006.

[3] O. Simeone, "A brief introduction to machine learning for engineers," *Found. Trends Signal Process.*, vol. 12, nos. 3–4, pp. 200–431, 2018. doi: 10.1561/2000000102.

[4] D. J. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[5] E. Angelino, M. J. Johnson, and R. P. Adams, "Patterns of scalable Bayesian inference," *Found. Trends Signal Process.*, vol. 9, nos. 2–3, pp. 119–247, Nov. 2016. doi: 10.1561/2200000052.

[6] J. Knoblauch, J. Jewson, and T. Damoulas, "Generalized variational inference," 2019, arXiv:1904.02063.

[7] P. D. Grünwald, I. J. Myung, and M. A. Pitt, *Advances in Minimum Description Length: Theory and Applications*. Cambridge, MA: MIT Press, 2005.

[8] A. A. Alemi, I. Fischer, J. V. Dillon, and K. Murphy, "Deep variational information bottleneck," 2016, arXiv:1612.00410.

[9] T. Zhang, "Information-theoretic upper and lower bounds for statistical estimation," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1307–1321, 2006.

[10] B. Guedj and L. Pujol, "Still no free lunches: The price to pay for tighter PAC-Bayes bounds," 2019, arXiv:1910.04460.

[11] M. I. Belghazi, A. Baratin, S. Rajeswar, S. Ozair, Y. Bengio, A. Courville, and R. D. Hjelm, "MINE: Mutual Information Neural Estimation," 2018, arXiv:1801.04062.

[12] A. Beck, *First-Order Methods in Optimization*, vol. 25. Philadelphia: SIAM, 2017.

[13] Y. Pantazis, D. Paul, M. Fasoulakis, Y. Stylianou, and M. Katsoulakis, "Cumulant GAN," 2020, arXiv:2006.06625.

Alexander Shapiro, Yao Xie, and Rui Zhang

On the Characteristic Rank for Matrix and Tensor Completion

In this lecture note, we discuss a fundamental concept, referred to as the *characteristic rank*, that suggests a general framework for characterizing the basic properties of various low-dimensional models used in signal processing. We illustrate this framework through two examples—matrix and three-way tensor completion problems—and consider basic properties, including the identifiability of matrices and tensors, given partial observations. We consider cases without observation noise to illustrate the principle.

Digital Object Identifier 10.1109/MSP.2020.3046233
Date of current version: 24 February 2021

Relevance

The characteristic rank provides a fundamental tool for determining the "order" of low-rank structures, such as the rank of low-rank matrices and the rank of three-way tensors. The concept of characteristic rank was introduced in [6], where it was used to establish necessary and sufficient conditions to determine the "recoverability" of low-rank matrices. The characteristic rank can also be generally applied to determine the "intrinsic" degrees of freedom in other low-rank manifold structures. Such instances include determining the number of hidden nodes in one-layer neural networks and establishing the number of sources in blind demixing problems, as shown in [7].

Prerequisites

To better comprehend the concepts discussed in this lecture note, readers are expected to have a good background in linear algebra, multivariate calculus, and basic concepts of measure theory, which we will explain whenever we run into them. Suggested references are

[4] and [3]. In the following, we review some basic concepts.

Manifold of low-rank matrices

Consider the set of $n_1 \times n_2$ matrices of rank r, denoted $\mathcal{M}_r$. Note that the rank is no larger than the dimension of the matrix: $r \leq \min\{n_1, n_2\}$. It is known that such a set of rank-r matrices $\mathcal{M}_r$ forms a smooth manifold in the space $\mathbb{R}^{n_1 \times n_2}$, and the dimension of the manifold is given by

$$\dim(\mathcal{M}_r) = r(n_1 + n_2 - r). \quad (1)$$

A matrix $A \in \mathcal{M}_r$ can be represented in the form $A = VW^\top$, where V and W are matrices of the respective order $n_1 \times r$ and $n_2 \times r$, both of full column rank r. Thus, we can view (V, W) as a parametrization of $\mathcal{M}_r$. Note that the number of involved parameters is $r(n_1 + n_2)$, which is larger than the dimension of $\mathcal{M}_r$; this is because V and W in the preceding representation are not unique.

Three-way tensor

Another example we will consider is the (three-way) tensor $X \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. It is said that X has rank one if $X = a \circ b \circ c$, where a, b, and c are vectors of the respective dimensions n_1, n_2, and n_3 and "$\circ$" denotes the vector outer product. That is, every element of tensor X can be written as the product $X_{ijk} = a_i b_j c_k$. The smallest number r such that tensor X can be represented as the sum r of rank-one tensors is called the *rank* of X. The corresponding decomposition is often referred to as the (*tensor*) *rank-decomposition* or the *canonical polyadic decomposition* [8], [10]. We would like to remark that our method can apply to higher-order tensors, as well.

Problem statement

Matrix completion

Let us start by considering the problem of reconstructing an $n_1 \times n_2$ matrix of a given rank r, while observing its entries $M_{ij}, (i,j) \in \Omega$, for an index set $\Omega \subset \{1,...,n_1\} \times \{1,...,n_2\}$ of cardinality $\mathsf{m} = |\Omega|$. This is known as the *exact matrix completion problem* [1], which is now well studied. The conditions for recovery have been derived assuming that entries are missing at random, and the performance guarantees are given in a probabilistic sense. Here, we aim to approach the problem from a geometric perspective, which can possibly lead to a deterministic and more intuitive answer. There are two basic difficulties associated with this problem, namely, the existence of the solution and the uniqueness of the solution; that is, whether such a matrix does exist and if so, whether it is unique. Fundamentally, these questions are related to the identifiability of low-rank matrices, which we define as follows.

Definition 1: Local identifiability of low-rank matrix completion problem

Let $Y \in \mathcal{M}_r$ be such that $[Y]_{ij} = M_{ij}$, $(i,j) \in \Omega$. [Thus, rank(Y) = r.] It is said that the matrix completion problem is locally identifiable at Y if there exists a neighborhood $\mathcal{N} \subset \mathbb{R}^{n_1 \times n_2}$ of Y such that for any $Y' \in \mathcal{N}$ and $Y' \neq Y$ with $[Y']_{ij} = M_{ij}, (i,j) \in \Omega$, the rank of Y' is different from r.

Uniqueness of tensor decomposition

Uniqueness is the key question related to tensor rank decomposition. Here, we consider the following tensor decomposition problem: given a three-way tensor X, we would like to find the associated matrix factors A, B, and C of the respective order $n_1 \times r, n_2 \times r$ and $n_3 \times r$ such that $X = [\![A, B, C]\!]$, meaning that $X = \Sigma_{i=1}^{r} a^i \circ b^i \circ c^i$, with a^i, b^i, and c^i being the ith columns of the respective matrices A, B, and C. Clearly, the decomposition $X = [\![A, B, C]\!]$ is invariant with respect to permutations of the rank-one components and rescaling of the columns of matrices A, B, and C by factors $\lambda_{1i}, \lambda_{2i}$, and λ_{3i} such that $\lambda_{1i}\lambda_{2i}\lambda_{3i} = 1, i = 1, \ldots, r$. We first introduce the global identifiability and the local identifiability of the tensor.

Definition 2: Global identifiability of tensor decomposition

The decomposition $X = [\![A, B, C]\!]$ (globally) identifies rank r if it is unique; i.e., if $X = [\![A', B', C']\!]$ is another decomposition of tensor X, with matrices A', B', C', being of the respective order $n_1 \times r', n_2 \times r', n_3 \times r'$ and $r' = r$, then both decompositions are the same up to the corresponding permutation and rescaling. It is said that the rank-r decomposition is generically identifiable if for almost every $(A, B, C) \in \mathbb{R}^{n_1 \times r} \times \mathbb{R}^{n_2 \times r} \times \mathbb{R}^{n_3 \times r}$, the corresponding tensor $X = [\![A, B, C]\!]$ identifies rank r.

Definition 3: Local identifiability of tensor decomposition

We say that $(A, B, C) \in \mathbb{R}^{n_1 \times r} \times \mathbb{R}^{n_2 \times r} \times \mathbb{R}^{n_3 \times r}$ is locally identifiable if there is a neighborhood $\mathcal{N}$ of (A, B, C) such that $(A', B', C') \in \mathcal{N}$ and $[\![A', B', C']\!] = [\![A, B, C]\!]$ imply that (A', B', C') can be obtained from (A, B, C) by the corresponding rescaling. We say that model (n_1, n_2, n_3, r) is generically locally identifiable if almost every $(A, B, C) \in \mathbb{R}^{n_1 \times r} \times \mathbb{R}^{n_2 \times r} \times \mathbb{R}^{n_3 \times r}$ is locally identifiable.

Like the matrix completion problem, it is also possible to consider a tensor completion problem: reconstructing a tensor of a given rank when only a subset of the entries is observed. The respective local and global identifiability concepts can be similarly defined.

Solutions

Matrix completion

Reparameterization of matrix completion problem

Let us start with the matrix completion problem by using the following parametrization. Consider the set $\mathcal{X}$ of $n_1 \times n_2$ matrices X such that $[X]_{ij} = 0, (i,j) \in \Omega$ (when adding such matrices to the solutions, the obtained matrices remain consistent with the observations). We can view $\mathcal{X}$ as a linear space of dimension $\dim(\mathcal{X}) = n_1 n_2 - \mathsf{m}$. Then, the matrix completion problem has a solution if and only if there exist respective matrices V and W of rank r and $X \in \mathcal{X}$ such that $[VW^\top + X]_{ij} = M_{ij}, (i, j) \in \Omega$. Let Θ be the set of vectors θ formed from the components of (V, W, X). Note that Θ is a subset of the vector space of dimension $r(n_1 + n_2) + n_1 n_2 - \mathsf{m}$.

Characteristic rank

The matrix completion parametrization can be considered as a mapping that assigns matrix $VW^{\top}+X$ to a vector of parameters $\theta=(V,W,X)\in\Theta$. With this mapping, we can define the so-called Jacobian matrix $\Delta(\theta)$, which is the matrix of the partial derivatives of $VW^{\top}+X$ with respect to components of vector θ. Then, we associate this mapping with its characteristic rank, defined as

$$\mathfrak{r}=\max_{\theta\in\Theta}\{\operatorname{rank}(\Delta(\theta))\}. \tag{2}$$

Note that the characteristic rank $\mathfrak{r}$ does not depend on the order in which the parameters are arranged.

The characteristic rank has the following properties: the rank of $\Delta(\theta)$ is equal to $\mathfrak{r}$ for almost every $\theta\in\Theta$. By *almost every* we mean that the set of such $\theta\in\Theta$ for which $\operatorname{rank}(\Delta(\theta))\neq\mathfrak{r}$ has a Lebesgue measure of zero. Moreover, the set $\{\theta\in\Theta:\operatorname{rank}(\Delta(\theta))=\mathfrak{r}\}$ forms an open subset of Θ. It follows that the rank of $\Delta(\theta)$ is constant and equals $\mathfrak{r}$ in a neighborhood of almost every $\theta\in\Theta$. This result implies that the characteristic rank is an intrinsic quantity associated with the "degrees of freedom" of the problem, regardless of the value of the parameters.

Implication of characteristic rank on matrix completion

We can also look at the characteristic rank from the following point of view. Consider the tangent space $\mathcal{T}_{\mathcal{M}_r}(Y)$ to the manifold $\mathcal{M}_r$ at the point $Y=VW^{\top}\in\mathcal{M}_r$. We have that

$$\operatorname{rank}(\Delta(\theta))=\dim(\mathcal{T}_{\mathcal{M}_r}(Y))+\dim(\mathcal{X})-\dim(\mathcal{T}_{\mathcal{M}_r}(Y)\cap\mathcal{X}). \tag{3}$$

The relation (3) can be explained as follows. Generically, the image of the considered mapping $VW^{\top}+X$ forms a smooth manifold in the image space, at least locally. The tangent space to this manifold at the considered point is the sum of the tangent space to $\mathcal{M}_r$ (from the parameterization $VW^{\top}$) and the linear space $\mathcal{X}$ in the image space. On the other hand, this tangent space is generated by columns of the Jacobian matrix $\Delta(\theta)$ (or, in other words, by the differential of the mapping), and its dimension is equal to the rank of $\Delta(\theta)$. Then, the right-hand side of (3) is the usual formula for the dimension of the sum of two linear spaces $\mathcal{T}_{M_r}(Y)$ and $\mathcal{X}$. Hence, from (3) and the definition of the characteristic rank (2), we have that

$$\mathfrak{r}=\dim(\mathcal{T}_{\mathcal{M}_r}(Y))+\dim(\mathcal{X})-\inf_{Y\in\mathcal{M}_r}\{\dim(\mathcal{T}_{\mathcal{M}_r}(Y)\cap\mathcal{X})\}. \tag{4}$$

By Sard's theorem [5], we have that the image of the set Θ by the mapping $\theta\mapsto VW^{\top}+X$ has the Lebesgue measure zero if and only if $\mathfrak{r}<n_1n_2$. That is, if $\mathfrak{r}<n_1n_2$, generically, the problem of reconstructing a matrix of rank r by observing its entries $M_{ij},(i,j)\in\Omega$ is unsolvable. By *generically* we mean that the set of rank-r solutions with components matching $M_{ij},(i,j)\in\Omega$ has the Lebesgue measure of zero in the corresponding vector space of dimension m.

In other words, if the characteristic rank is smaller than the dimension n_1n_2 of the image space, any solution of rank r is unstable: this means that arbitrarily small changes of the data values M_{ij} make the rank-r solution unattainable. Note that the characteristic rank is a function of the index set Ω and does not depend on the observed values M_{ij}. In particular, because of (4) we have that $\mathfrak{r}<n_1n_2$ if $\mathsf{m}>r(n_1+n_2-r)$. For example, if $n_1=n_2=10$, $r=3$, we have $\mathfrak{r}<100$ if $\mathsf{m}>3\times(10+10-3)=51$. Since the characteristic rank is the dimension of the image of the mapping, if it is smaller than the dimension n_1n_2 of the image space, then it is "thin," i.e., of measure zero in the image space.

Well-posedness condition

By the preceding discussion, we have that if

$$\mathcal{T}_{\mathcal{M}_r}(Y)\cap\mathcal{X}=\{0\}, \tag{5}$$

at least for one point $Y\in\mathcal{M}_r$, then

$$\mathfrak{r}=\dim(\mathcal{T}_{\mathcal{M}_r}(Y))+\dim(\mathcal{X}). \tag{6}$$

Conversely, if (6) holds, condition (5) is satisfied for all $Y\in\mathcal{M}_r$ except for a set of measure zero in $\mathcal{M}_r$. Condition (5) implies local identifiability at Y. Generically, the matrix completion problem is locally identifiable if and only if condition (6) holds, which is referred to as the well-posedness condition in [6]. Figure 1 illustrates the point. Generically, the intersection of $\mathcal{T}_{M_r}(Y)$ and $\mathcal{X}$ gives the tangent space to the intersection of M_r and $\mathcal{X}$. When the intersection of $\mathcal{T}_{M_r}(Y)$ and $\mathcal{X}$ is $\{0\}$, we have well-posedness and local uniqueness.

Simple example

Here, we illustrate the characteristic rank using a simple example of a two-by-two rank-one matrix $M=vw^{\top}$, with partial observations at $\Omega=\{(1,1),(2,2)\}$. Then,

$$X=\begin{bmatrix}0 & x_{12}\\ x_{21} & 0\end{bmatrix},$$

$\theta=(v_1,v_2,w_1,w_2,x_{12},x_{21})$. We have

$$\Delta(\theta)=\frac{\partial(vw^{\top}+X)}{\partial\theta}=\begin{bmatrix}w_1 & 0 & v_1 & 0 & 0 & 0\\ w_2 & 0 & 0 & v_1 & 1 & 0\\ 0 & w_1 & v_2 & 0 & 0 & 1\\ 0 & w_2 & 0 & v_2 & 0 & 0\end{bmatrix}.$$

It can be verified that $\operatorname{rank}(\Delta(\theta))=4$ for almost every $\theta\in\Theta$; thus, $\mathfrak{r}=4$. Consider a possible rank-one solution to this problem. The tangent space of the rank-one manifold $\dim(\mathcal{T}(\mathcal{M}_r))=2+2-1=3$, and $\dim(\mathcal{X})=2$; $\mathfrak{r}<\dim(T(M_r))+\dim(\mathcal{X})$, and the well-posedness condition (6)

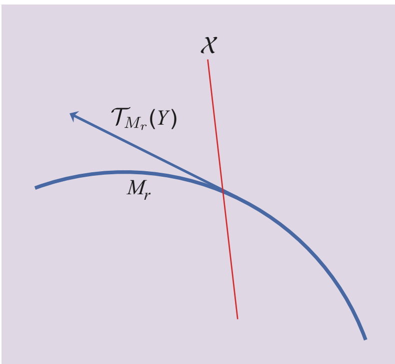

FIGURE 1. The well-posedness condition for the matrix completion problem.

is not satisfied. Indeed, the rank-one solution to this problem is not unique: it can be any $x_{12}x_{21} = c$, where c is the product of the observed diagonal elements.

On the other hand, if $\Omega = \{(1,1), (1,2), (2,1)\}$,

$$X = \begin{bmatrix} 0 & 0 \\ 0 & x_{22} \end{bmatrix},$$

$\theta = (v_1, v_2, w_1, w_2, x_{22})$. We have

$$\Delta(\theta) = \frac{\partial(vw^\top + X)}{\partial\theta} = \begin{bmatrix} w_1 & 0 & v_1 & 0 & 0 \\ w_2 & 0 & 0 & v_1 & 0 \\ 0 & w_1 & v_2 & 0 & 0 \\ 0 & w_2 & 0 & v_2 & 1 \end{bmatrix}.$$

It can be verified that $\operatorname{rank}(\Delta(\theta)) = 4$ for almost every $\theta \in \Theta$, and thus $\mathfrak{r} = 4$. The rank of the tangent space is $2 + 2 - 1 = 3$; the dimension of $\mathcal{X}$ is one. Thus, $\mathfrak{r} = \dim(\mathcal{T}(\mathcal{M}_r)) + \dim(\mathcal{X})$, and the well-posedness condition (6) is satisfied. Indeed, the solution to this matrix completion problem is unique.

Checking conditions

Although the preceding simple example is easy to check, evaluating the characteristic rank in a closed form is not always easy for larger instances. Nevertheless, the rank of the Jacobian matrix can be computed numerically, and hence condition (6) can be verified for a considered index set Ω and rank r. Clearly, local identifiability is a necessary condition for global identifiability (i.e., for the global uniqueness of the solution). There is an example in [11] that shows that local identifiability does not imply global identifiability. Assuming that all observed entries are different than zero, necessary and sufficient conditions for global identifiability are known when $r = 1$. Those conditions are the same for local identifiability (see [6] for more details). Giving necessary and sufficient conditions for global identifiability for a general r and Ω could be too difficult and out of reach. On the other hand, the simple dimensionality condition (6) gives a verifiable condition, at least for local identifiability.

Tensor decomposition

Invoking characteristic rank on three-way tensor

Here, we briefly discuss local identifiability for tensor decomposition. For three-way tensor recovery, we can consider the mapping

$$\mathcal{G}_r : (A, B, C) \mapsto [\![A, B, C]\!]. \tag{7}$$

Similar to (2), the characteristic rank $\mathfrak{r}$ of the mapping is given by the maximal rank of its Jacobian matrix, and it has generic properties similar to the ones discussed for the matrix completion problem. Note that $\mathfrak{r}$ is always less than or equal to $r(n_1 + n_2 + n_3 - 2)$. This follows by counting the number of elements in (A, B, C) and making corrections for the scaling factors.

The model (n_1, n_2, n_3, r) is generically locally identifiable if and only if the following condition for the characteristic rank holds:

$$\mathfrak{r} = r(n_1 + n_2 + n_3 - 2). \tag{8}$$

The condition (8) is necessary for generic global identification and can be numerically verified by computing the rank of the Jacobian matrix of the mapping $\mathcal{G}_r$. Let us note that in a similar spirit, it is also possible to give conditions for the local identifiability of the tensor completion problem when only a set of observed values of the tensor components is available (i.e., tensor completion problems). To do so, we need to set up an appropriate mapping and study the associated characteristic rank. We can refer to [2, Sec. 3.2] and the references therein for a discussion of the uniqueness (identifiability) of tensor rank decompositions. For the tensor completion problem, local identifiability does not imply global identifiability, even in the rank-one case (e.g., [9]).

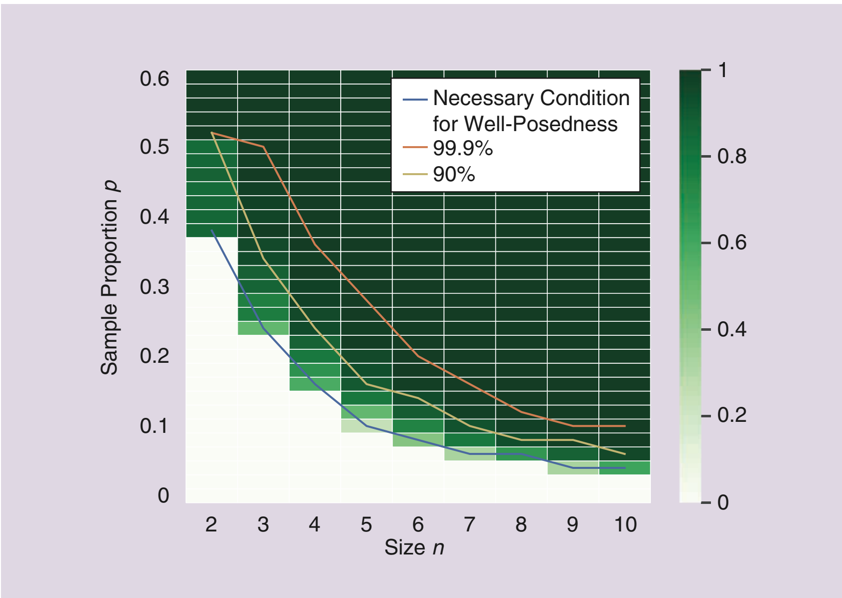

FIGURE 2. Recovering a three-way tensor with missing data: the probability of well-posedness being satisfied versus a theoretical prediction. The blue line corresponds to $p = (3n-2)/n^3$; the yellow and the orange lines correspond to the sampling proportion whereby the well-posedness condition is empirically satisfied with a probability of 90% and 99.9%.

Computational example

Here, we present a numerical example to illustrate how to use the characteristic rank to study a three-way tensor's completion problem. Consider the case where the tensor entries are randomly sampled. Assume the size of each dimension of the tensor is n, and thus the size of the tensor is $\mathbb{R}^{n \times n \times n}$. The proportion of the observed entries is p, and the total number of observed entries is $\mathsf{m} = \lceil pn^3 \rceil$, where $\lceil x \rceil$ is the ceiling function for rounding up to the nearest integer. For each p, we randomly choose m observations from the tensor. For the reported experiments, we used $n = 2, \ldots, 10$ and $p = 0.02, 0.04, \ldots, 0.6$. To validate the theoretical results, we perform 300 random trials for each combination of n and p. For each trial, we generate a, b, $c \in \mathbb{R}^n$ such that the entries of the vector are independent identically distributed normal random variables with zero mean and unit variance. With a, b, and c and index set Ω, the Jacobian matrix of $a \circ b \circ c + X$ can be computed, where $[X]_{ijk} = 0, (i,j,k) \in \Omega$. If the rank of the Jacobian equals $3n - 2 + n^3 - m$, we conclude that the well-posedness condition is satisfied for the instance [see (6)].

Finally, we report the proportion of the random instances satisfying the condition. As mentioned, the necessary condition for well-posedness is that $\mathsf{m} \geq 3n - 2$ [see (8)]. This requires, approximately, $p \geq (3n-2)/n^3$. Figure 2 gives the probability that well-posedness is satisfied for rank-one tensors under different tensor sizes and sampling proportions. Note that the empirical results match the theoretical prediction well. Moreover, it can be observed that as the tensor size becomes large, the well-posedness condition is satisfied with a small sampling proportion.

What we have learned

In this lecture note, we explained how to use a fundamental concept, namely, the characteristic rank, to answer essential questions, such as identifiability, when given observations of a low-rank structure (e.g., low-rank matrices and low-rank three-way tensors). The framework involved a few steps. We first found the map that associated the truth to the observations, then studied the Jacobian matrix of the map to find the characteristic rank and compared the characteristic rank with respective conditions that were problem specific (such as the well-posedness condition). Once the concepts are understood, the analysis usually involves only basic multivariate calculus. The benefit is that the tool can generally be applicable to study other problems with low-rank structures. We have considered cases without observation noise to illustrate the principle. When there are additive Gaussian noises, statistical goodness-of-fit tests can be developed based on the framework [7].

Acknowledgments

This work was funded, in part, by National Science Foundation CAREER Awards CCF-1650913, CMMI-2015787, DMS-1938106, and DMS-1830210.

Authors

Alexander Shapiro (ashapiro@isye.gatech.edu) received his Ph.D. degree from the Ben-Gurion University. He is the A. Russell Chandler III Chair of and a professor in the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA. He is an author of more than 140 research articles in peer-reviewed journals and a coauthor of several books, and he has served on numerous editorial boards, including as editor-in-chief of *Mathematical Programming, Series A*. His research interests are stochastic programming, risk analysis, and simulation-based optimization. In 2013, he received the Khachiyan Prize for Lifetime Accomplishments in Optimization, and in 2018, he was awarded the Dantzig Prize by the Mathematical Optimization Society and Society for Industrial and Applied Mathematics. In 2020 he was elected to the National Academy of Engineering.

Yao Xie (yao.xie@isye.gatech.edu) received her Ph.D. degree in electrical engineering from Stanford University. She is an associate professor and the Harold R. and Mary Anne Nash Early Career Professor in the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA, where she is also an associate director of the Machine Learning Center. She is an associate editor of *IEEE Transactions on Signal Processing*, *Sequential Analysis: Design Methods and Applications*, and *INFORMS Journal on Data Science*, and she serves on the editorial board of *Journal of Machine Learning Research*. Her research interests include statistics (in particular, sequential analysis and sequential change-point detection), machine learning, and signal processing. She received the National Science Foundation CAREER Award in 2017. He is a Member of IEEE.

Rui Zhang (ruizhang_ray@gatech.edu) received his M.S. degree in statistics from the University of Michigan, in 2017. He is a doctoral student in the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30332, USA. His research interests include change-point detection, machine learning, and statistics.

References

[1] E. J. Candès and T. Tao, “The power of convex relaxation: Near-optimal matrix completion,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, 2010. doi: 10.1109/TIT.2010.2044061.

[2] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009. doi: 10.1137/07070111X.

[3] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, vol. 71. Philadelphia: SIAM, 2000.

[4] W. Rudin, *Principles of Mathematical Analysis*, vol. 3. New York: McGraw-Hill, 1964.

[5] A. Sards, “The measure of critical values of differential maps,” *Bull. Amer. Math. Soc.*, vol. 48, no. 12, pp. 883–890, 1942. doi: 10.1090/S0002-9904-1942-07811-6.

[6] A. Shapiro, Y. Xie, and R. Zhang, “Matrix completion with deterministic pattern: A geometric perspective,” *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 1088–1103, 2019. doi: 10.1109/TSP.2018.2885494.

[7] A. Shapiro, Y. Xie, and R. Zhang, “Goodness-of-fit tests on manifolds,” 2019, arXiv:1909.05229.

[8] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning,” *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, 2017. doi: 10.1109/TSP.2017.2690524.

[9] M. Singh, A. Shapiro, and R. Zhang, “Rank one tensor completion problem,” 2020, arXiv:2009.10533.

[10] L. Sorber, M. V. Barel, and L. D. Lathauwer, “Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-(L_r,L_r,1) terms, and a new generalization,” *SIAM J. Optim.*, vol. 23, no. 2, pp. 695–720, 2013. doi: 10.1137/120868323.

[11] E. B. Wilson and J. Worcester, “The resolution of six tests into three general factors,” *Proc. Nat. Acad. Sci. U.S.A*, vol. 25, no. 2, pp. 73–77, 1939. doi: 10.1073/pnas.25.2.73.

SP