Novel Tensor GSVD Predicting Ovarian Cancer Survival and Response to Platinum-Based Chemotherapy

Theodore E. Schomay,[1,2] Katherine A. Aiello,[1,2] and Orly Alter[1,2,3]

[1]Scientific Computing and Imaging (SCI) Institute, [2]Department of Bioengineering, and [3]Department of Human Genetics

University of Utah, Salt Lake City, UT

Motivation

The number of large-scale high-dimensional datasets recording different aspects of a single phenomenon is growing, accompanied by a need for frameworks that can create a single coherent model from multiple order-matched tensors, with one-to-one mappings among the columns across all but one of the corresponding dimensions among the tensors. The recent higher-order generalized singular value decomposition (HO GSVD) is the only simultaneous decomposition to date of more than two such datasets that is by definition exact, and which mathematical properties allow interpreting its variables and operations in terms of the similar as well as dissimilar among the datasets.[1] The HO GSVD, however, is limited to datasets arranged in matrices. It builds upon the two-matrix GSVD, which we previously reformulated as a comparative spectral decomposition.[2,3]

We recently demonstrated a novel tensor GSVD in the comparative modeling of two such datasets.[4,5] Rather than simplifying the complex structure of the datasets to a single matrix, the tensor GSVD uses it to simultaneously find the similarities and dissimilarities, i.e., patterns of varying relative significance, between the two tensors. Here, we define and prove the mathematical properties of the tensor GSVD.

Fig. 1. Our recent tensor GSVD comparison of two patient- and platform-matched genomic datasets uncovered patterns of DNA alterations that are correlated with ovarian cancer survival and response to chemotherapy.[4]

Biological Results

Our recent comparison of the genomes of ovarian tumor and normal cells from the same set of patients, measured by two different methods, uncovered several chromosome arm-wide patterns of DNA copy-number alterations that are correlated with a patient's survival and response to platinum-based chemotherapy.[4] To uncover these patterns of variation across the patients, that are consistent across the measurements, but exclusive to the tumor relative to the normal genome, we used the tensor GSVD. For >30 years prior, the best predictor of ovarian cancer survival was the tumor's stage; ~25% of primary ovarian tumors are resistant to platinum, yet no diagnostic existed to distinguish resistant from sensitive tumors before the treatment.

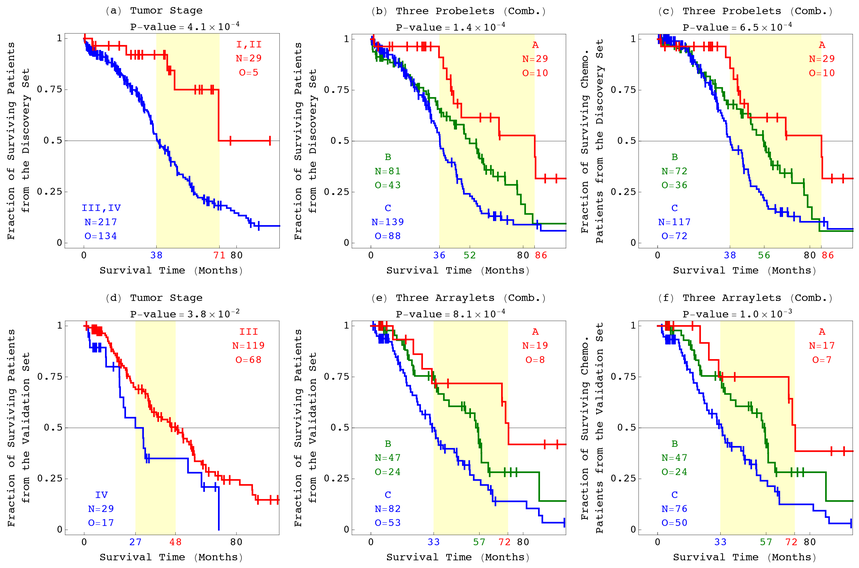

Fig. 2. Survival analyses of the discovery (a-c) and validation (d-f) sets of patients classified by (a,d) tumor stage at diagnosis, and (b,c,e,f) the tensor GSVD.

Mathematical Framework: Tensor GSVD

Definition and Construction. We define the tensor GSVD, which simultaneously factorizes the two, e.g., third-order, tensors $\mathcal{D}_i \in \mathbb{R}^{K_i \times L \times M}$, $i = 1, 2$, as

$$\mathcal{D}_i = \mathcal{R}_i \times_a U_i \times_b V_x \times_c V_y.$$

The matrices V_x and V_y, identical in both factorizations, and U_i are obtained from the GSVDs of the tensors unfolded to preserve the x-, y-, and z-axes, respectively. The tensors are now superpositions of corresponding pairs of rank-one "subtensors" $\mathcal{S}_i(a,b,c) = u_{i,a} \otimes v_{x,b}^T \otimes v_{y,c}^T$, i.e.,

$$\mathcal{D}_i = \sum_{a=1}^{LM} \sum_{b=1}^{L} \sum_{c=1}^{M} \mathcal{R}_{i,abc} \mathcal{S}_i(a,b,c).$$

We prove that the tensor GSVD (*i*) always exists, (*ii*) has the same uniqueness properties as the GSVD, and (*iii*) extends the GSVD and the higher-order singular value decomposition from decompositions of two matrices or one tensor, respectively, to a decomposition of two tensors.[6]

Interpretation. Interpreting the tensor GSVD as a framework for comparatively modeling two datasets, we define the significance of the subtensor $\mathcal{S}_1(a,b,c)$ in $\mathcal{D}_1$ relative to that of $\mathcal{S}_2(a,b,c)$ in $\mathcal{D}_2$ as a function of the ratio of the corresponding superposition coefficients $\mathcal{R}_{1,abc}/\mathcal{R}_{2,abc}$. We prove that this ratio is equivalent to that of the generalized singular values of $\mathcal{D}_1$ and $\mathcal{D}_2$ unfolded to preserve the z-axis.

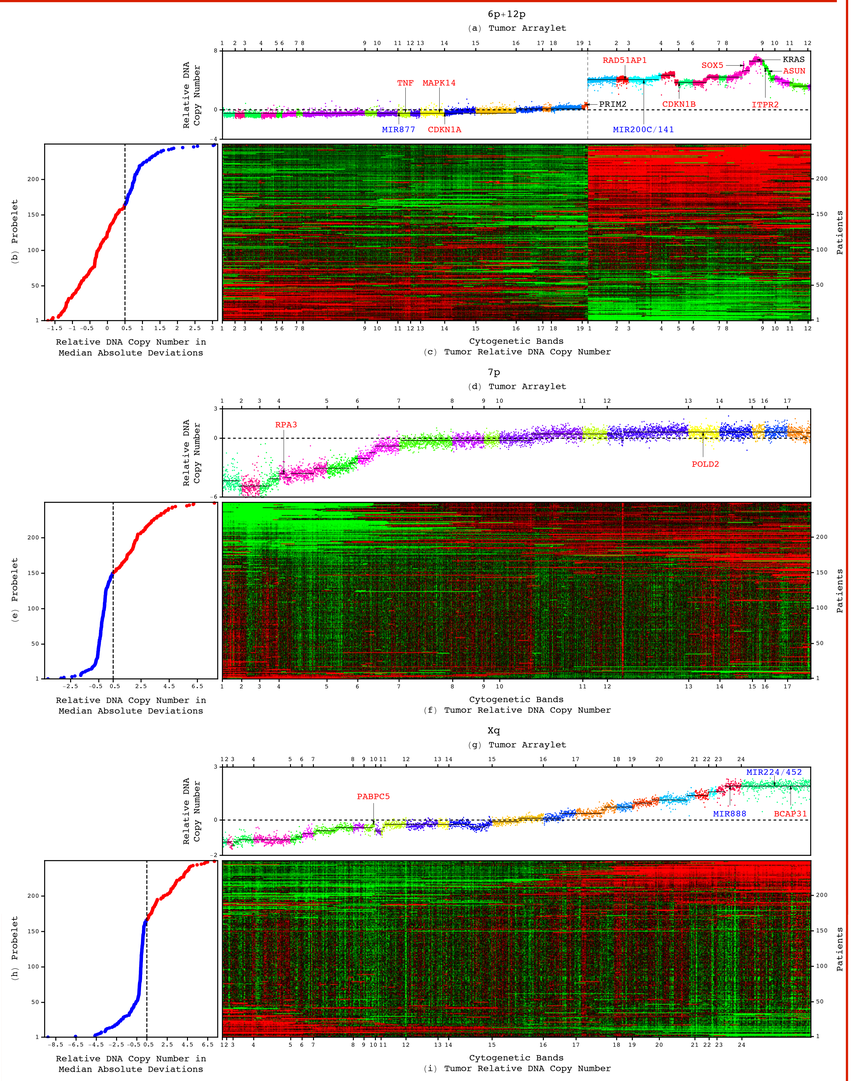

Fig. 3. The tensor GSVD of patient- and platform-matched ovarian tumor and normal genomic profiles uncovered (a,d,g) previously unknown chromosome arm-wide signatures that (b,e,h) classify the patients into two groups of significantly different prognoses, and are mathematically and biologically exclusive to (c,f,i) the tumor profiles.

References

1. S. P. Ponnapalli, M. A. Saunders, C. F. Van Loan, O. Alter, "A Higher-Order Generalized Singular Value Decomposition for Comparison of Global mRNA Expression from Multiple Organisms," *PLoS One* 6, e28072 (2011)
2. O. Alter, P. O. Brown, D. Botstein, "Generalized Singular Value Decomposition for Comparative Analysis of Genome-Scale Expression Datasets of Two Different Organisms," *PNAS USA* 100 (6), 3351–3356 (2003)
3. C. H. Lee, B. O. Alpert, P. Sankaranarayanan, O. Alter, "GSVD Comparison of Patient-Matched Normal and Tumor aCGH Profiles Reveals Global Copy-Number Alterations Predicting Glioblastoma Multiforme Survival," PLoS One 7, e30098 (2012); http://dx.doi.org/10.1371/journal.pone.0030098
4. P. Sankaranarayanan,* T. E. Schomay,* K. A. Aiello, O. Alter, "Tensor GSVD of Patient- and Platform-Matched Tumor and Normal DNA Copy-Number Profiles Uncovers Chromosome Arm-Wide Patterns of Tumor-Exclusive Platform-Consistent Alterations Encoding for Cell Transformation and Predicting Ovarian Cancer Survival," *PLoS One* 10, e0121396 (2015); http://dx.doi.org/10.1371/journal.pone.0121396
5. T. E. Schomay, K. A. Aiello and O. Alter, "The Tensor GSVD: A Comparative Spectral Decomposition for Two Column-Matched and Row-Independent Tensors," *2016 Workshop on Tensor Decompositions and Applications* (Leuven, Belgium, January 18–22, 2016).
6. T. E. Schomay, K. A. Aiello, O. Alter, *in preparation.*

This research is supported by National Cancer Institute (NCI) U01 Grant CA-202144 (http://physics.cancer.gov/network/UniversityofUtah.aspx), the Utah Science, Technology and Research (USTAR) Initiative, National Human Genome Research Institute (NHGRI) R01 Grant HG-004302 and National Science Foundation (NSF) CAREER Award DMS-0847173.