# 5 Tensor Decompositions

Tensors can be defined using three equivalent ways:

1. Equivalence classes in a free vector space.

2. Transformation properties upon a change of vector space basis.

3. Linearizer of multilinear maps.

A variation of definition 1 is taught in quantum mechanics, and definition 2 is usually taught in relativity. Tensors are also used heavily in condensed matter theory nowadays to approximate a wave function as a tensor network. A tensor network consists of nodes that are tensors and edges that represent a contraction of covariant and contravariant indices between connected tensors.

In this Chapter, we will discuss all these ideas and learn how to use tensors to represent data and perform eigen-decompositions of tensors.

## 5.1 Tensor Product Space from a Free Vector Space Generated by Vector Spaces

Let $V_1, \ldots, V_k$ be vector spaces over $\mathbb{R}$ of dimension $n_1, \ldots, n_k$, respectively. Let $\mathcal{F}$ be the free vector space over $\mathbb{R}$ generated by the Cartesian product $V_1 \times V_2 \times \ldots \times V_k$; i.e., $\mathcal{F}$ consists of elements that are finite $\mathbb{R}$-linear combinations of elements in $V_1 \times V_2 \times \ldots \times V_k$, treating the distinct elements of $V_1 \times V_2 \times \ldots \times V_k$ as distinct basis elements of $\mathcal{F}$. Let $\mathcal{M}$ be a subspace in $\mathcal{F}$ generated by elements of the form: $\forall x_j, y_j \in V_j$, $j = 1, \ldots, k$, and $\alpha \in \mathbb{R}$

$$(x_1, \ldots, x_{i-1}, x_i + y_i, x_{i+1}, \ldots, x_k) - (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k) - (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_k)$$

and

$$\alpha(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k) - (x_1, \ldots, x_{i-1}, \alpha x_i, x_{i+1}, \ldots, x_k)$$

for all $i = 1, \ldots, k$.

**Definition 5.1** (Tensor Product Space)**.** *The tensor product $V_1 \otimes \cdots \otimes V_k$ of vector spaces $V_1, \ldots, V_k$ is the quotient space $\mathcal{F}/\mathcal{M}$.*

**REMARK 5.1.** *You can think of the quotient as imposing the following equivalence relations on $\mathcal{F}$:*

$$(x_1, \ldots, x_{i-1}, x_i + y_i, x_{i+1}, \ldots, x_k) \sim (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k) + (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_k)$$

*and*

$$\alpha(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_k) \sim (x_1, \ldots, x_{i-1}, \alpha x_i, x_{i+1}, \ldots, x_k).$$

**EXERCISE 5.1.** *If $\{e_j^{(i)}\}_{j=1}^{n_i}$ is a basis of $V_i$, then show that $\{e_{j_1}^{(1)} \otimes \cdots \otimes e_{j_k}^{(k)}\}_{j_1=1,\ldots,j_k=1}^{n_1,\ldots,n_k}$, which denote the equivalence classes $\{[(e_{j_1}^{(1)}, \cdots, e_{j_k}^{(k)})]\}_{j_1=1,\ldots,j_k=1}^{n_1,\ldots,n_k}$ in $\mathcal{F}/\mathcal{M}$, is a basis of the tensor product $\otimes_{i=1}^k V_i$. Thus, the dimension of $\otimes_{i=1}^k V_i$ as a vector space is $\prod_{i=1}^k n_i$.*

**Definition 5.2** ($(p,q)$-tensors)**.** *Let $V_1, \ldots, V_p, W_1, \ldots, W_q$ be vector spaces. The elements of the tensor product space*

$$T^p_q = V_1 \otimes \cdots \otimes V_p \otimes W_1^* \otimes \cdots \otimes W_q^*$$

*are called order-$(p,q)$ tensors. When all $V_i$ and $W_j$ are equal to a common vector space $V$, we denote the above tensor product space as $T^p_q(V)$ and call its elements order-$(p,q)$ tensors on $V$.*

**REMARK 5.2.** *In the basis of $T^p_q$ described in Exercise 5.1, we can represent a tensor $t$ in terms of its components $t^{i_1 \cdots i_p}_{j_1 \cdots j_q}$. In the literature, the notation $t$ and its component representation $t^{i_1 \cdots i_p}_{j_1 \cdots j_q}$ are often used interchangeably.*

## 5.2 Transformation of Tensors

### 5.2.1 Change of Basis

Let $V$ be a vector space with a basis $\{e_1, \ldots, e_n\}$. Let $V^*$ be the dual vector space of $V$ with a dual basis $\{e^{*1}, \ldots, e^{*n}\}$ satisfying $e^{*j}(e_i) = \delta^j_i$. Consider changing the basis of $V$ to $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ which is related to the original basis via

$$\boxed{\tilde{e}_k = \sum_{i=1}^n A^i{}_k e_i}, \tag{5.1}$$

where $A$ is an invertible matrix such that

$$e_i = \sum_{k=1}^n (A^{-1})^k{}_i \tilde{e}_k. \tag{5.2}$$

Then, under the change of basis described in (5.1), the dual basis needs to satisfy

$$\tilde{e}^{*\ell}(\tilde{e}_k) = \delta^\ell_k = \tilde{e}^{*\ell}\left(\sum_{i=1}^n A^i{}_k e_i\right) = \sum_{i=1}^n A^i{}_k \tilde{e}^{*\ell}(e_i) \Rightarrow \tilde{e}^{*\ell}(e_i) = (A^{-1})^\ell{}_i.$$

Hence,

$$\boxed{\tilde{e}^{*\ell} = \sum_{j=1}^n (A^{-1})^\ell{}_j e^{*j}}.$$

That is, the transformation matrix for the dual basis is the inverse transpose of that for the basis.

Now, for any vector $v \in V$, we need

$$v = \sum_{i=1}^n v^i e_i = \sum_{k=1}^n \tilde{v}^k \tilde{e}_k. \tag{5.3}$$

Using (5.2), this equality yields:

$$\sum_{i=1}^{n}\sum_{k=1}^{n} v^i {(A^{-1})^k}_i \tilde{e}_k = \sum_{k=1}^{n} \tilde{v}^k \tilde{e}_k \Rightarrow \boxed{\tilde{v}^k = \sum_{i=1}^{n} {(A^{-1})^k}_i \, v^i}.$$

Thus, the vector coefficients transform like the dual basis elements.

Similarly, any element $w \in V^*$ can be written as

$$w = \sum_{j=1}^{n} w_j \, e^{*j} = \sum_{\ell=1}^{n} \tilde{w}_\ell \, \tilde{e}^{*\ell},$$

which takes exactly the same form as (5.3), except that the superscripts and subscripts are exchanged. Thus, since the dual basis elements $\tilde{e}^{*\ell}$ transform like the vector coefficients $\tilde{v}^k$, the dual vector coefficients $\tilde{w}_\ell$ should transform like the basis elements in order to satisfy the transformation equality; that is, we must have

$$\boxed{\tilde{w}_\ell = \sum_{j=1}^{n} {A^j}_\ell w_j}.$$

To summarize, the dual vector coefficients $\tilde{w}_\ell$ transform like the basis elements and are said to be covariant, because they co-vary with the basis elements. By contrast, the vector coefficients $\tilde{v}^k$ transform like the dual basis elements and are said to be contravariant.

**Definition 5.3** (Einstein Summation Convention)**.** *Repeated covariant and contravariant indices are assumed to be summed over all possible values. For example, $v^i e_i$ means $\sum_i v^i e_i$.*

### 5.2.2 Transformation of Tensors

Using Exercise 5.1, we can now define a $(p,q)$-tensor $t \in T^p_q(V)$ on vector space $V$ as a set of coefficients with $p$ contravariant indices and $q$ covariant indices. That is

$$\tilde{t}^{k_1\cdots k_p}_{\ell_1\cdots\ell_q} = {(A^{-1})^{k_1}}_{i_1} \cdots {(A^{-1})^{k_p}}_{i_p} {A^{j_1}}_{\ell_1} \cdots {A^{j_q}}_{\ell_q} \, t^{i_1\cdots i_p}_{j_1\cdots j_q}.$$

For simplicity, we have assumed that $V_i = W_i = V$, but a similar transformation formula holds for a general $(p,q)$-tensor.

In particular, matrices are tensors that transform in a specific way. To see this, let $V$ and $W$ be real vector spaces with a basis $\{e_1,\ldots,e_n\}$ and $\{f_1,\ldots,f_m\}$, respectively. We have seen in Chapter 1 that a linear map $M : V \to W$ can be expressed in a matrix form in these bases as:

$$M(e_j) = \sum_{i=1}^{m} {M^i}_j f_i.$$

**REMARK 5.3.** *Note that we now put the row index as a superscript and the column index as a subscript. It will become clear why this convention makes sense shortly.*

Suppose we change the bases to $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ and $\{\tilde{f}_1, \ldots, \tilde{f}_m\}$:

$$\tilde{e}_k = \sum_{j=1}^{n} A^j{}_k e_j \text{ and } \tilde{f}_\ell = \sum_{i=1}^{m} B^i{}_\ell f_i,$$

where $A$ and $B$ are invertible matrices such that

$$e_j = \sum_{k=1}^{n} \left(A^{-1}\right)^k{}_j \tilde{e}_k \text{ and } f_i = \sum_{\ell=1}^{m} \left(B^{-1}\right)^\ell{}_i \tilde{f}_\ell.$$

In these new bases, the action of $M$ is

$$\begin{aligned} M(\tilde{e}_k) = M(\sum_{j=1}^{n} A^j{}_k e_j) &= \sum_{j=1}^{n} A^j{}_k M(e_j) = \sum_{j=1}^{n} A^j{}_k \left(\sum_{i=1}^{m} M^i{}_j f_i\right) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} M^i{}_j A^j{}_k f_i = \sum_{\ell=1}^{m} \left(\sum_{i=1}^{m}\sum_{j=1}^{n} \left(B^{-1}\right)^\ell{}_i M^i{}_j A^j{}_k\right) \tilde{f}_\ell. \end{aligned}$$

Hence,

$$\tilde{M}^\ell{}_k = \sum_{i=1}^{m}\sum_{j=1}^{n} \left(B^{-1}\right)^\ell{}_i A^j{}_k M^i{}_j.$$

$M : V \to W$ transforms like a $(1,1)$-tensor and is an element of $W \otimes V^*$.

### 5.2.3 Tensor Product Space as a Vector Space

Above, we wrote $M \in W \otimes V^*$ and not $M \in V^* \otimes W$. The two are actually equivalent. To understand this subtle concept, let us recall that a tensor product space is a vector space. Hence, given two vector spaces $V_1$ and $V_2$ of dimension $n_1$ and $n_2$, respectively, and elements $a \in V_1$ and $b \in V_2$, we should be able to express $a \otimes b$ as an $n_1 n_2 \times 1$ column vector. Let $a = (a^1, \ldots, a^{n_1})^t$ in some basis of $V_1$. The convention is

$$a \otimes b = \begin{pmatrix} a^1 b \\ a^2 b \\ \vdots \\ a^{n_1} b \end{pmatrix}. \tag{5.4}$$

Now, consider the linear maps $A : V_1 \to W_1$ and $B : V_2 \to W_2$. We can extend the pair to a linear map, denoted $A \otimes B$, from $V_1 \otimes V_2$ to $W_1 \otimes W_2$ as follows:

$$\forall a \in V_1, b \in V_2,\ (A \otimes B)(a \otimes b) = (Aa) \otimes (Bb). \tag{5.5}$$

Then, the matrix representation of $A \otimes B$ consistent with (5.4) and (5.5) is the following block matrix

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B & \cdots & A_{1n_1}B \\ A_{21}B & A_{22}B & \cdots & A_{2n_1}B \\ \vdots & \vdots & & \vdots \\ A_{m_11}B & A_{m_12}B & \cdots & A_{m_1n_1}B \end{pmatrix},$$

where $m_1$ is the dimension of $W_1$.

**REMARK 5.4.** *Check this claim.*

This product, known as the Kronecker product, holds for any two matrices $A$ and $B$. In particular, if $\{e_1, \dots, e_{n_1}\}$ and $\{f_1, \dots, f_{n_2}\}$ are the standard bases of $V$ and $W$, respectively, then

$$f_i \otimes e^{*j} = f_i(e_j)^t\,,$$

which is 1 at the $i$-th row and $j$-th column, and 0 elsewhere; we have identified the dual vector with its transpose. Hence,

$$M^i{}_j\, f_i \otimes e^{*j} = M,$$

justifying our choice $M \in W \otimes V^*$. But, we also have the Kronecker product

$$e^{*j} \otimes f_i = f_i(e_j)^t\,,$$

implying that $M$ can be also viewed as an element of $V^* \otimes W$. This isomorphism $W \otimes V^* \cong V^* \otimes W$ allows us not to care too much about the specific ordering of vector spaces in the tensor product space.

## 5.3 Multilinear Maps

**Definition 5.4** (Multilinear Maps)**.** *Let $V_1, \dots, V_k, W$ be vector spaces. A map $L : V_1 \times \cdots \times V_k \to W$ is called multi-linear if it is linear in each argument when other remaining arguments are fixed. The set of all multilinear maps from $V_1 \times \cdots V_k$ to $W$ is denoted as $\mathscr{L}(V_1, \dots, V_k; W)$.*

**Example 5.1.** *Let $\{e_j^{(i)}\}_{j=1}^{\dim V_i}$ denote a basis of $V_i$, and $W = V_1 \otimes \cdots \otimes V_k$. Any vector $v^{(i)} \in V_i$ can be written as $v^{(i)} = \sum_j v^j e_j^{(i)}$. The inclusion map*

$$\imath : V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k,$$

*defined by*

$$(v^{(1)}, \dots, v^{(k)}) \mapsto v^{(1)} \otimes \cdots \otimes v^{(k)},$$

*is multilinear.*

**EXERCISE 5.2.** *Show that a tensor $t \in V_1 \otimes \cdots \otimes V_p \otimes W_1^* \otimes \cdots \otimes W_q^*$ can be also viewed as a multilinear map $t \in \mathscr{L}(V_1^*, \dots, V_p^*, W_1, \dots, W_q; \mathbb{R})$.*

**EXERCISE 5.3.** *Let $U, V, W$ be vector spaces. Let $\mathrm{Hom}(V, W)$ denote the set of all linear maps from $V$ to $W$. Show that*

$$\mathscr{L}(U, V; W) \cong \mathrm{Hom}(U, \mathrm{Hom}(V, W)) \cong \mathrm{Hom}(V, \mathrm{Hom}(U, W)).$$

**Theorem 5.1.** *For any multilinear map $L \in \mathscr{L}(V_1, \ldots, V_k; W)$, there exists a unique linear map $\varphi : V_1 \otimes \cdots \otimes V_k \to W$ such that the below diagram commutes*

$$\begin{tikzcd} V_1 \times \cdots V_k \arrow[r, "\imath"] \arrow[dr, "L"'] & V_1 \otimes \cdots \otimes V_k \arrow[d, "\varphi"] \\ & W \end{tikzcd}$$

*i.e. $L = \varphi \circ \imath$, where $\imath$ is the inclusion map.*

*Proof.* Because of the multilinearity of $L$, $L = \varphi \circ \imath$ if we define

$$\varphi(e^{(1)}_{j_1} \otimes \cdots \otimes e^{(k)}_{j_k}) = L(e^{(1)}_{j_1}, \ldots, e^{(k)}_{j_k})$$

for all Cartesian products of basis elements and extend its action via linearity to $V_1 \otimes \cdots \otimes V_k$. If $\varphi'$ is another linearizer, then $\varphi - \varphi' \equiv 0$ on $V_1 \otimes \cdots \otimes V_k$, implying that $\varphi = \varphi'$ as linear maps. $\square$

**REMARK 5.5.** *The tensor product space $V_1 \otimes \cdots \otimes V_k$ is thus called the universal linearizer of $\mathscr{L}(V_1, \ldots, V_k; W)$.*

## 5.4 Tensor Products, Tensor Rank and Contractions

A tensor product space $T^p_q = V_1 \otimes \cdots \otimes V_p \otimes W^*_1 \otimes \cdots \times W^*_q$ is a vector space, so we can take a tensor product of tensor product spaces. That is, given another tensor product space $T^r_s = \tilde{V}_1 \otimes \cdots \otimes \tilde{V}_r \otimes \tilde{W}^*_1 \otimes \cdots \times \tilde{W}^*_s$, we can define a multiplication of tensors $t \in T^p_q$ and $s \in T^r_s$ as

$$\begin{aligned} T^p_q \times T^r_s &\to T^p_q \otimes T^r_s \\ (t, s) &\mapsto t \otimes s. \end{aligned}$$

**REMARK 5.6.** *By the bilinearity of tensor product operation, the components of $t \otimes s$ are just products of the components of $t$ and $s$.*

**REMARK 5.7.** *Upon shuffling the vector spaces, we can see that $T^p_q \otimes T^r_s \cong T^{p+r}_{q+s}$.*

**REMARK 5.8.** *Not every element in $T^p_q \otimes T^r_s$ can be rewritten as a tensor product of some $t \in T^p_q$ and $s \in T^r_s$. Those that can be written as $t \otimes s$ are said to be decomposed as a product of lower order tensors. This is the first example of tensor decomposition.*

**REMARK 5.9.** *The tensor product operation allows us to define a tensor algebra on the direct sum*

$$T(V) = \bigoplus_{p=0}^{\infty} T^p(V).$$

*This is the starting point of homological and exterior algebra in mathematics.*

**Definition 5.5** (Tensor Rank). *Let $t \in T^p = V_1 \otimes \cdots \otimes V_p$. The rank of $t$ is*

$$\operatorname{rank}(t) = \arg\min_r \left\{ r \in \mathbb{N} \;\middle|\; t = \sum_{i=1}^{r} v_{(i)}^{(1)} \otimes \cdots \otimes v_{(i)}^{(p)}, \text{ where } v_{(i)}^{(j)} \in V_j \right\}.$$

**REMARK 5.10.** *A rank-1 tensor of order $(p, 0)$ is thus a tensor product of $p$ vectors.*

**REMARK 5.11.** *The rank of a tensor is the minimum number of rank-1 tensors that generate the tensor, and the resulting decomposition is called the canonical polyadic (CP) decomposition. This definition is analogous to the definition of matrix rank, which is the smallest number $k$ such that the matrix can be written as a sum of $k$ outer products of vectors:*

$$M = \sum_{i=1}^{k} u_i v_i^t.$$

**REMARK 5.12.** *Computing the rank of a generic tensor is an NP-hard problem. An exact CP decomposition is thus also NP-hard.*

**Definition 5.6** (Kruskal Notation). *Let $\lambda \in \mathbb{R}^k$, and let $U^{(n)} \in \mathbb{R}^{I_n \times k}$, for $i = 1, \dots, p$, denote an $I_n \times k$ matrix with unit $\ell_2$-norm columns. Then,*

$$[\![\lambda; U^{(1)}, \dots, U^{(p)}]\!] \equiv \sum_{j=1}^{k} \lambda_j \, u_{(j)}^{(1)} \otimes \cdots \otimes u_{(j)}^{(p)},$$

*where $u_{(j)}^{(i)}$ denotes the $j$-th column of $U^{(i)}$.*

Using this notation, we can phrase the low-rank approximation problem for tensors as follows: given fixed $k \in \mathbb{N}$ and $\mathcal{Y} \in V_1 \otimes \cdots \otimes V_p$, where $\dim(V_n) = I_n$, find $\lambda \in \mathbb{R}^k, U^{(n)} \in \mathbb{R}^{I_n \times k}$, such that $\mathcal{Y} \approx \mathcal{X} \equiv [\![\lambda; U^{(1)}, \dots, U^{(p)}]\!]$. In order to define what $\approx$ means, we need a notion of distance on the tensor product space. We will accomplish this task in Section 5.9.

**Definition 5.7** (Contraction). *Let $T_q^p = V_1 \otimes \cdots \otimes V_p \otimes W_1^* \otimes \cdots \times W_q^*$. Let $v^{(i)} \in V_i$ and $\varphi^{(j)} \in W_j^*$. The linear map $C_s^r : T_q^p \to T_{q-1}^{p-1}$, for $1 \le r \le p$ and $1 \le s \le q$ where $V_r = W_s$, defined by*

$$C_s^r(v^{(1)} \otimes \cdots \otimes v^{(p)} \otimes \varphi^{(1)} \otimes \cdots \otimes \varphi^{(q)}) = \varphi^{(s)}(v^{(r)}) v^{(1)} \otimes \cdots v^{(\hat{r})} \cdots \otimes v^{(p)} \otimes \varphi^{(1)} \otimes \cdots \varphi^{(\hat{s})} \cdots \otimes \varphi^{(q)}$$

*is called a contraction map.*

**REMARK 5.13.** *By linearity of the contraction map, we only need to examine the action of $C_s^r$ on the basis elements $e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_p}^{(p)} \otimes f^{*(1)\,j_1} \otimes \cdots \otimes f^{*(q)\,j_q}$. In component notation, the condition $f^{*(s)\,j_s}(e_{i_r}^{(r)}) = \delta_{i_r}^{j_s}$ implies that the matching contravariant and covariant indices are summed under the contraction map.*

**Example 5.2.** *In general relativity, a contraction of the Riemann curvature tensor $R^i{}_{jk\ell}$ yields the Ricci tensor $R_{j\ell} = R^k{}_{jk\ell}$.*

**Example 5.3.** *Let $U, V, W$ be vector spaces, and $M : V \to W$ a linear map. We have seen previously that $M \in W \otimes V^*$. Given any tensor $t \in U \otimes V$, we can take the tensor product $t \otimes M \in U \otimes V \otimes W \otimes V^*$ and then contract the second and fourth indices to get*

$$s^{ik} = t^{ij} M^k_{\ j}\,,$$

*which define the components of a tensor $s \in U \otimes W$. Equivalently, we can convert a tensor in $U \otimes V$ into a tensor in $U \otimes W$ via the linear map $I \otimes M : U \otimes V \to U \otimes W$, where $\forall u \in U, v \in V$, $(I \otimes M)(u \otimes v) = u \otimes Mv$.*

## 5.5 Tensor Embedding of Vectorial Data (Tensorization)

Tensor product arises naturally in quantum mechanics as a way of decomposing a wave function into independent degrees of freedom or approximating the wave function of a many-body system. Recent advances in condensed matter theory has led to the idea of representing a data point $x \in \mathbb{R}^n$ as a tensor $t^{i_1 \cdots i_p}$. Then, a model can take a weighted sum of this tensor as

$$\hat{y} = w_{i_i \cdots i_p} t^{i_1 \cdots i_p}$$

and use the value as a predictor in classification. Notice that the weights $w_{i_i \cdots i_p}$ themselves are tensors, and we need ways to tune these parameters. Often, the weight tensor has too many components, and we need simplifying assumptions. These assumptions can arise in the form of approximating the weights using low-rank tensor decompositions; e.g. density matrix renormalization group (DMRG) utilizes matrix product states to approximate low-energy states. We will not discuss these approaches in this course.

Another way of using tensors that represent data would be to treat a tensor product space as a vector space and use known linear techniques on the vector space:

**Example 5.4.** *Consider the concentric annulus data shown in Figure 5.1. The tensor product embedding described by*

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos \frac{\pi x}{4} \\ \sin \frac{\pi x}{4} \end{pmatrix} \otimes \begin{pmatrix} \cos \frac{\pi y}{4} \\ \sin \frac{\pi y}{4} \end{pmatrix} \in \mathbb{R}^2 \otimes \mathbb{R}^2$$

*can linearly separate the three clusters. Viewing $\mathbb{R}^2 \otimes \mathbb{R}^2$ as $\mathbb{R}^4$, Figure 5.1 demonstrates this separation in PC2 and PC3 directions. Intuitively, this embedding maps $x \sim 0$ to a "spin-0" state $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $x \sim \pm 2$ to "spin-1" states $\begin{pmatrix} 0 \\ \pm 1 \end{pmatrix}$, and similarly for $y$.*

**Example 5.5.** *A time series $(x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$, where $N = I + J - 1$ for some positive*

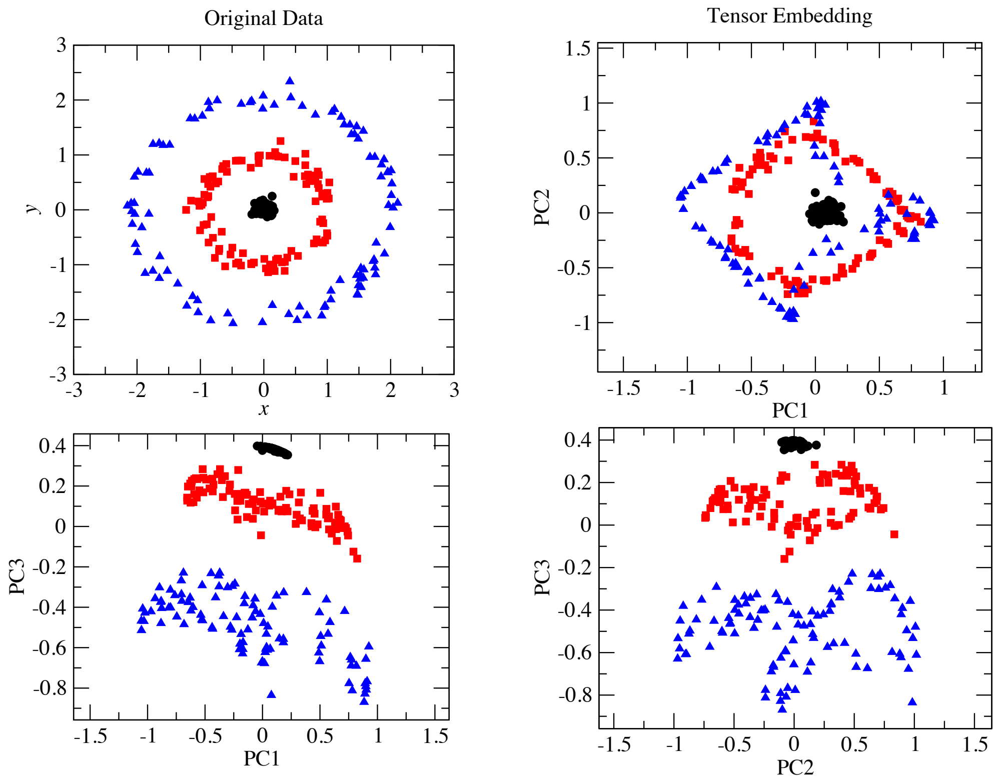


Figure 5.1: Tensor embedding of original concentric annulus data. The embedded data can be separated by linear hyperplanes in PC2 and PC3 directions.

*integers $I$ and $J$, can be transformed into the following Hankel matrix*

$$H = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_J \\ x_2 & x_3 & x_4 & \cdots & x_{J+1} \\ x_3 & x_4 & x_5 & \cdots & x_{J+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_I & x_{I+1} & x_{I+2} & \cdots & x_N \end{pmatrix}$$

*which can be decomposed via SVD into interpretable components; see Figure 5.2. This transformation is known as Hankelization, and the signal decomposition method is called the singular spectrum analysis. Hankelizing each column or row of the above Hankel matrix will yield an order-3 Hankel tensor. Decomposing this tensor may inform useful properties of the original time series signal.*

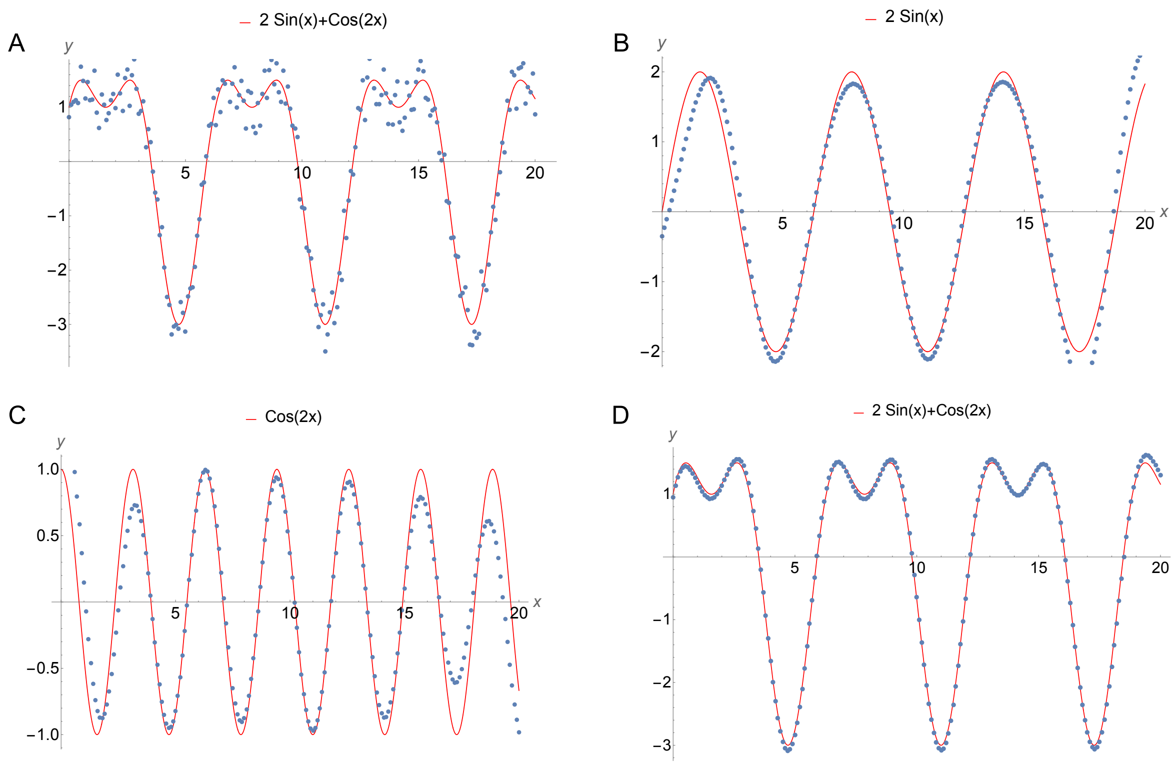


Figure 5.2: Singular Spectrum Analysis of 1D signal. (A) Random noise uniformly distributed in $[-0.5, 0.5]$ was added to the true signal $2\sin x + \cos 2x$. The noisy data were sampled at 201 points located at $x = 0.1k$, for $k = 0, \ldots, 200$. (B-D) The sampled 1D vector was Hankelized into a $100 \times 102$ Hankel matrix $H$, and SVD was performed on $H$. In (B), Hankelization of the first two terms in SVD by averaging the skew-diagonal terms yielded a 1D signal reconstructing the component $2\sin x$. In (C), Hankelization of the next two terms in SVD yielded a 1D signal reconstructing the component $\cos 2x$. Summing the two components from (B) and (C) reconstructed the full denoised signal $2\sin x + \cos 2x$, as shown in (D).

## 5.6 Symmetric and Alternating Tensors

Let $\pi_p$ be the symmetric group of permutations on $p$ elements. Then, $\forall \sigma \in \pi_p, t = t^{i_1,\ldots,i_p} e_{i_1} \otimes \cdots \otimes e_{i_p}$, we can define an action of $\sigma$ on $T^p(V)$ as

$$\sigma(t) = t^{i_{\sigma(1)},\ldots,i_{\sigma(p)}} e_{i_1} \otimes \cdots \otimes e_{i_p}.$$

**Definition 5.8** (Symmetric Tensor). *A tensor $t \in T^p(V)$ is called symmetric if $\forall \sigma \in \pi_p$, $\sigma(t) = t$. The set of all symmetric tensors of order $(p, 0)$ is denoted as $\mathrm{Sym}^p(V)$ or $S^p(V)$.*

**Definition 5.9** (Alternating Tensor). *A tensor $t \in T^p(V)$ is called alternating if $\forall \sigma \in \pi_p$, $\sigma(t) = \mathrm{sign}(\sigma) t$, where $\mathrm{sign}(\sigma) = 1$ if $\sigma$ is an even permutation and $\mathrm{sign}(\sigma) = -1$ if $\sigma$ is an odd permutation. The set of all alternating tensors of order $(p, 0)$ is denoted as $\bigwedge^p(V)$.*

**REMARK 5.14.** *Let $\{e_j\}$ and $\{f_i\}$ be bases of $V$ and $W$. We noted above that a matrix $M : V \to W$ is a $(1, 1)$-tensor; specifically, $M \in W \otimes V^*$. Recall that the transpose of*

$M$ *is* $M^t : W^* \to V^*$, *i.e.* $M^t \in V^* \otimes W$. *In component notation,* $M^t = (M^t)_j{}^i \, e^{*j} \otimes f_i$, *where, by definition,* $(M^t)_j{}^i = M^i{}_j$. *When we say* $M$ *is symmetric and write* $M = M^t$, *we mean that* $(M^t)_j{}^i = M^j{}_i$ *upon identifying* $V \cong V^*$ *and* $W \cong W^*$. *Thus, if* $M$ *is symmetric,* $M^i{}_j \equiv (M^t)_j{}^i = M^j{}_i$.

## 5.7 Tensor Unfolding

Current tensor decomposition algorithms rely on reorganizing the tensor components into a matrix. There is no unique way of "matricizing" a tensor, so we first need to introduce some jargon to explain the techniques.

**Definition 5.10** (Mode)**.** *Let* $t \in V_1 \otimes \cdots \otimes V_p$. *In component notation, the set of all* $n$*-th index* $i_n$ *in* $t^{i_1 \cdots i_p}$ *is called the* $n$*-th mode.*

**Definition 5.11** (Fiber)**.** *A mode-*$n$ *fiber of* $t^{i_1 \cdots i_p}$ *is an array obtained by fixing all indices* $i_\ell$, *for* $\ell \neq n$, *and varying the* $n$*-th index* $i_n$.

**REMARK 5.15.** *A mode-1 fiber of a matrix is thus a column vector, and a mode-2 fiber of a matrix is a row vector.*

**Definition 5.12** (Mode-$n$ Unfolding (a.k.a. matricization or flattening))**.** *Let* $t \in V_1 \otimes \cdots \otimes V_p$, *where* $\dim(V_i) = I_i$. *Then, the mode-*$n$ *unfolding of* $t$ *is a matrix* $\mathbf{t}_{(n)}$ *of dimension* $I_n \times (\prod_{i \neq n} I_i)$ *obtained by organizing the mode-*$n$ *fibers of* $t$ *along the columns. The tensor component* $t^{i_1 \cdots i_p}$ *is mapped to the* $i_n$*-th row and* $j$*-th column of* $\mathbf{t}_{(n)}$, *where*

$$j = 1 + \sum_{\ell=1, \ell \neq n}^{p} \left[ (i_\ell - 1) J_\ell \right] ,$$

*where*

$$J_\ell = \begin{cases} \displaystyle\prod_{m=\ell+1}^{n-1} I_m, & \ell \leq n-2 \\ 1, & \ell = n-1 \\ \displaystyle\prod_{r=1}^{n-1} I_r \prod_{s=\ell+1}^{p} I_s, & \ell \geq n+1 \end{cases}$$

**REMARK 5.16.** *It does not matter how we organize the columns as long as we are consistent in subsequent formulae.*

**Example 5.6.** *Let* $\mathcal{X} \in T^2$ *with components*

$$\mathcal{X} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

*Then,*

$$\mathbf{X}_{(1)} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \mathcal{X} \quad \textit{and} \quad \mathbf{X}_{(2)} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = \mathcal{X}^t.$$

**Example 5.7.** *Let $\mathcal{X} = v_1 \otimes v_2 \otimes \cdots \otimes v_p$, where $v_i$ are vectors. Then,*

$$\mathbf{X}_{(n)} = v_n(v_{n+1} \otimes \cdots v_p \otimes v_1 \otimes \cdots v_{n-1})^t,$$

*where $v_n$ is represented as a column vector and $v_{n+1} \otimes \cdots v_p \otimes v_1 \otimes \cdots v_{n-1}$ is the Kronecker product of remaining vectors.*

**Definition 5.13** ($n$-mode Product)**.** *Let $\mathcal{X} \in V_1 \otimes \cdots \otimes V_p$, and let $A \in W \otimes V_n^*$ represent a linear map from $V_n$ to $W$. Then, $\mathcal{Y} \equiv \mathcal{X} \times_n A = C_1^n(\mathcal{X} \otimes A)$ is an order-$(p, 0)$ tensor with components*

$$\mathcal{Y}^{i_1 \ldots, i_{n-1} j i_{n+1} \ldots i_p} = X^{i_1 \ldots i_p} A^j{}_{i_n}.$$

**EXERCISE 5.4.** *In terms of unfolded matrices, show that*

$$\mathcal{Y} = \mathcal{X} \times_n A \iff \mathbf{Y}_{(n)} = \mathbf{A}\mathbf{X}_{(n)}.$$

*Show that*

$$\mathcal{X} \times_m A \times_n B = \mathcal{X} \times_n B \times_m A, \ \textit{for } m \neq n$$

*and*

$$\mathcal{X} \times_n A \times_n B = \mathcal{X} \times_n (BA), \ \textit{where } A : V_n \to W_1, \ B : W_1 \to W_2.$$

**EXERCISE 5.5.** *Let $\mathcal{X}$ be a tensor of order $(p, 0)$. Let $\mathcal{Y} = \mathcal{X} \times_1 U_1 \times_2 U_2 \times_3 \cdots \times_p U_p$. Then,*

$$\mathbf{Y}_{(n)} = U_n \mathbf{X}_{(n)} (U_{n+1} \otimes \cdots \otimes U_p \otimes U_1 \otimes \cdots \otimes U_{n-1})^t.$$

**REMARK 5.17.** *The SVD of a matrix $M$ can be written as $M = \Sigma \times_1 U \times_2 V$. The mode expansions are $M_{(1)} = U\Sigma V^t$ and $M_{(2)} = V \Sigma U^t = M^t$.*

## 5.8 Frobenius Inner Product and Norm of Tensors

Similar to our discussion of low-rank approximations of matrices, we need a notion of distance on the space of tensors in order to judge whether an approximation is good or not. We also need to generalize the notion of orthogonality to generalize SVD to tensors. For these purposes, we can directly generalize the definitions of Frobenius inner product and norm of matrices, Definition 1.38 and Definition 1.39, to tensors:

**Definition 5.14** (Frobenius Inner Product)**.** *The map $\langle \cdot, \cdot \rangle_F : T^p(V) \times T^p(V) \to \mathbb{R}$ defined by $\langle t, s \rangle_F = \sum_{i_1, \ldots, i_p} t^{i_1 \ldots i_p} s^{i_1 \ldots i_p}$ is called the Frobenius inner product.*

**Definition 5.15** (Frobenius Norm)**.** *The map $\| \cdot \|_F : T^p(V) \to \mathbb{R}$ defined by $\|t\|_F = \sqrt{\langle t, t \rangle_F}$ is called the Frobenius norm of tensor $t$.*

**REMARK 5.18.** *We can similarly define the Frobenius inner product and norm on $T^p_q$.*

**REMARK 5.19.** *If $\mathbf{t}_{(n)}$ is the mode-$n$ unfolding of tensor $t$, then $\|t\|_F = \|\mathbf{t}_{(n)}\|_F$.*

## 5.9 CP Approximation

**Definition 5.16** (Khatri-Rao Product). *Let $A$ and $B$ be $I \times K$ and $J \times K$ matrices, respectively. Then, the Khatri-Rao product of $A$ and $B$, denoted $A \odot B$, is defined as*

$$A \odot B = (A_{:,1} \otimes B_{:,1} \ \cdots \ A_{:,K} \otimes B_{:,K}),$$

*where $\otimes$ represents the Kronecker product of two vectors.*

**Definition 5.17** (Hadamard Product). *Let $A$ and $B$ be $I \times J$ matrices. Then, the Hadamard product of $A$ and $B$, denoted $A * B$, is the element-wise product of the two matrices:*

$$(A * B)_{ij} = A_{ij} B_{ij}.$$

The Khatri-Rao product satisfies key algebraic properties outlined in Appendix A.8.

Let $V_n$ be a vector space of dimension $I_n$, for $n = 1, \ldots, p$. As previously mentioned, given a generic tensor $\mathcal{Y} \in V_1 \otimes \cdots \otimes V_p$, finding its exact CP decomposition, which may not even be essentially unique, is an NP-hard problem. As a computational compromise, we assume that the tensor rank of $\mathcal{Y}$ is not too large and attempt to find a low-rank tensor $\mathcal{X} = [\![\lambda; U^{(1)}, \ldots, U^{(p)}]\!]$ that has at most rank $K$, for some fixed $K$, and that best approximates $\mathcal{Y}$ w.r.t. the Frobenius norm. Here, $U^{(n)}$ is an $I_n \times K$ matrix with unit $\ell_2$-norm columns that are not necessarily orthogonal, and $\lambda \in \mathbb{R}^K$. We can try to solve this problem via an iterative algorithm. Namely, suppose we fix all $U$ matrices except for $U^{(n)}$, and then minimize $\|\mathcal{Y} - \mathcal{X}\|_F$ over $U^{(n)}$. Let $\mathbf{Y}_{(n)}$ denote the mode-$n$ unfolding of $\mathcal{Y}$. Using the result of Example 5.7, the mode-$n$ unfolding of $\mathcal{X}$ can be expressed as

$$\mathbf{X}_{(n)} = U^{(n)} \Lambda (U^{(n+1)} \odot \cdots \odot U^{(p)} \odot U^{(1)} \odot \cdots \odot U^{(n-1)})^t,$$

where $\Lambda = \mathrm{diag}(\lambda)$. Hence, we want to solve

$$\min_A \|\mathbf{Y}_{(n)} - A(U^{(n+1)} \odot \cdots \odot U^{(p)} \odot U^{(1)} \odot \cdots \odot U^{(n-1)})^t\|_F^2, \tag{5.6}$$

where $A = U^{(n)}\Lambda$ and iterate over $n$. To simplify notation, let $B = (U^{(n+1)} \odot \cdots \odot U^{(p)} \odot U^{(1)} \odot \cdots \odot U^{(n-1)})$. Then, setting the derivative of (5.6) with respect to $A$ to 0 yields

$$A(B^t B) = \mathbf{Y}_{(n)} B \Rightarrow A = \mathbf{Y}_{(n)} B (B^t B)^+, \tag{5.7}$$

where $^+$ denotes the Moore-Penrose pseudo-inverse. Theorem A.11 now implies that

$$A = \mathbf{Y}_{(n)} (U^{(n+1)} \odot \cdots \odot U^{(p)} \odot U^{(1)} \odot \cdots \odot U^{(n-1)})^{+^t}.$$

$\lambda$ is the vector of $\ell_2$-norm of the columns of $A$, and $U^{(n)}$ is retrieved from $A$ by normalizing the columns to have unit $\ell_2$-norm.

---

`CP Decomposition Pseudocode`

`Given` $\mathcal{Y} \in V_1 \otimes \cdots \otimes V_p$, $\dim V_n = I_n$.
**function** `CP-Approximation(`$\mathcal{Y}$`, rank` $K$`):`

```
Initialize U^(n) ∈ R^{I_n×K}, n = 1,...,p, with unit ℓ_2-norm columns
While fit improves and maximum iteration not reached:
        for n = 1,...,p:
            B = (U^(n+1) ⊙ ··· ⊙ U^(p) ⊙ U^(1) ⊙ ··· ⊙ U^(n-1))
            A = Y_(n) B(B^t B)^+
            λ = column norms of A.
            U^(n) = normalized columns of A.
Return λ, U^(1),...,U^(p).
```

---

**REMARK 5.20.** **Warning**: *Like many iterative algorithms, this algorithm is not guaranteed to converge to a global minimum solution. In some cases, the objective function may not even have a minimum, but only an infimum; in such a case, some rank-1 components in the decomposition may keep increasing without bound, while the objective function keeps decreasing very slowly.*

**REMARK 5.21.** *For order-3 tensors, Kruskal proved in 1977 that a CP solution to the optimization problem is essentially unique if*

$$2K + 2 \leq k_1 + k_2 + k_3,$$

*where $k_i$ is the Kruskal rank of $U^{(i)}$, defined as the greatest integer $r$ such that every set of $r$ columns of $U^{(i)}$ is linearly independent. Kruskal's sufficient condition is also necessary for $K = 2$ and $K = 3$, but not for $K > 3$.*

**REMARK 5.22.** *More recent results show that as the dimension of individual vector space $V_n$ increases, a generic low-rank tensor tends to have an essentially unique CP decomposition.*

## 5.10 Higher-Order SVD (HOSVD)

**Theorem 5.2** (HOSVD). *Let $V_n$ be a vector space of dimension $I_n$, for $n = 1, \ldots, p$. A tensor $\mathcal{X} \in V_1 \otimes \cdots \otimes V_p \equiv T^p$ can be decomposed as*

$$\mathcal{X} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \cdots \times_p U^{(p)} \tag{5.8}$$

*where*

1. *$U^{(n)} = (U^{(n)}_{(1)} \cdots U^{(n)}_{(I_n)})$ is an $I_n \times I_n$ orthogonal matrix, and*
2. *the core tensor $\mathcal{S} \in T^p$ is "all-orthogonal" and ordered; that is, the order-$(p-1, 0)$ subtensors $S^{i_n=\alpha}$ obtained by fixing the $n$-th index $i_n$ satisfy*
   (a) *All-orthogonality: $\forall n \in \{1, \ldots, p\}, \langle S^{i_n=\alpha}, S^{i_n=\beta}\rangle_F = 0$ for $\alpha \neq \beta$, and*
   (b) *Ordering: $\|S^{i_n=1}\|_F \geq \cdots \geq \|S^{i_n=I_n}\|_F \geq 0$.*

**REMARK 5.23.** *The core tensor $\mathcal{S}$ plays the same role as the diagonal matrix $\Sigma$ of singular values in matrix SVD. The $\Sigma$ matrix clearly satisfies the all-orthogonality and ordering properties.*

*Proof.* Consider the matrix SVD of the mode-$n$ unfolding $\mathbf{X}_{(n)}$ of $\mathcal{X}$:

$$\mathbf{X}_{(n)} = U^{(n)}\Sigma^{(n)}{V^{(n)}}^t, \tag{5.9}$$

where $\Sigma^{(n)} = \mathrm{diag}(\sigma_1^{(n)}, \dots, \sigma_{I_n}^{(n)})$, with $\sigma_1^{(n)} \geq \sigma_2^{(n)} \cdots \geq \sigma_{I_n}^{(n)} \geq 0$. Then, define

$$S = \mathcal{X} \times_1 {U^{(1)}}^t \times_2 {U^{(2)}}^t \cdots \times_p {U^{(p)}}^t.$$

By the orthogonality of $U^{(n)}$ matrices, we see that

$$\mathcal{X} = S \times_1 U^{(1)} \times_2 U^{(2)} \cdots \times_p U^{(p)}.$$

We claim that $S$ is the desired core tensor.

The mode-$n$ unfolding of $S$ is

$$\begin{aligned}\mathbf{S}_{(n)} &= {U^{(n)}}^t \mathbf{X}_{(n)} \left({U^{(n+1)}}^t \otimes \cdots \otimes {U^{(p)}}^t \otimes {U^{(1)}}^t \otimes \cdots \otimes {U^{(n-1)}}^t\right)^t \\ &= {U^{(n)}}^t \mathbf{X}_{(n)} \left(U^{(n+1)} \otimes \cdots \otimes U^{(p)} \otimes U^{(1)} \otimes \cdots \otimes U^{(n-1)}\right),\end{aligned}$$

where we have applied Theorem A.13. Substituting the expression (5.9), we get

$$\mathbf{S}_{(n)} = \Sigma^{(n)}{V^{(n)}}^t \left(U^{(n+1)} \otimes \cdots \otimes U^{(p)} \otimes U^{(1)} \otimes \cdots \otimes U^{(n-1)}\right).$$

All-orthogonality of $S^{i_n}$ is equivalent to the row-orthogonality of $\mathbf{S}_{(n)}$, which follows from the fact that the rows of ${V^{(n)}}^t$ are orthogonal and that $\left(U^{(n+1)} \otimes \cdots \otimes U^{(p)} \otimes U^{(1)} \otimes \cdots \otimes U^{(n-1)}\right)$ is an orthogonal transformation (Theorem A.15) that preserves the orthogonality of the rows of ${V^{(n)}}^t$. Furthermore, the Frobenius norm of the subtensor $S^{i_n=\alpha}$ is the Frobenius norm of the $\alpha$-th row of $\mathbf{S}_{(n)}$, which is equal to $\sigma_\alpha^{(n)}$. Hence, the ordering property is satisfied. □

**REMARK 5.24.** *In practice, we only need to compute the $I_n$-dimensional left singular vectors of $\mathbf{X}_{(n)}$ and not bother with computing the right singular vectors, which have a dimension $\prod_{k\neq n} I_k$ that is typically much larger then $I_n$. Then, compute $S = \mathcal{X} \times_1 {U^{(1)}}^t \cdots \times_p {U^{(p)}}^t$.*

**REMARK 5.25.** *Because $\mathcal{X}$ is obtained from $S \otimes U^{(1)} \otimes \cdots \otimes U^{(p)}$ by contracting all contravariant indices of $S$ with covariant indices of $U^{(n)}$, we can rewrite (5.8) as*

$$\mathcal{X} = \sum_{i_1,\dots,i_p} S^{i_1\dots i_p} U^{(1)}_{(i_1)} \otimes \cdots \otimes U^{(p)}_{(i_p)}, \tag{5.10}$$

*which is a tensor version of the matrix SVD, $M = \sum_i \sigma_i u_i v_i^t$, decomposing $\mathcal{X}$ as a sum of orthonormal rank-1 tensors.*

**REMARK 5.26.** *The subtensor $S^{i_n=\alpha}$ will be identically 0 if its corresponding $\sigma_\alpha^{(n)}$ is 0, since $\|S^{i_n=\alpha}\|_F = \sigma_\alpha^{(n)} = 0 \Rightarrow S^{i_n=\alpha} = 0$. $S$ is thus a compressed version of $\mathcal{X}$.*

**REMARK 5.27.** *Because the matrices $U^{(n)}$ are orthogonal, $\|\mathcal{X}\|_F^2 = \|\mathcal{S}\|_F^2 = \sum_{i=1}^{R_1}(\sigma_i^{(1)})^2 = \cdots = \sum_{i=1}^{R_p}(\sigma_i^{(p)})^2$, where $R_n$, called the n-mode rank or n-rank, is the matrix rank of $\mathbf{X}_{(n)}$.*

**Theorem 5.3** (Truncated HOSVD)**.** *Let $\hat{\mathcal{X}}$ be a truncation of $\mathcal{X}$ obtained by discarding the first few smallest-singular-value left singular vectors in HOSVD. That is, choose a threshold $C$, and rank the singular values $\{\sigma_1^{(1)}, \dots \sigma_{R_p}^{(p)}\}$; then, find the cutoff indices $I'_n$ such that $\sigma_{i_n}^{(n)} < C$ for $i_n > I'_n$. Set the components of $S$ containing indices $i_n > I'_n$ to 0 in (5.10) to obtain $\hat{\mathcal{X}}$. Then,*

$$\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \leq \sum_{i_1=I'_1+1}^{R_1} (\sigma_{i_1}^{(1)})^2 + \cdots + \sum_{i_p=I'_p+1}^{R_p} (\sigma_{i_p}^{(p)})^2.$$

*Proof.* Exercise. □

**REMARK 5.28.** *Unlike in the matrix case, the truncated approximation in Theorem 5.3 may not be the best possible approximation that satisfies the n-mode rank constraints. However, if the remaining singular values are small, say compared to the norm of $\mathcal{X}$, then $\hat{\mathcal{X}}$ still provides a good approximation of $\mathcal{X}$.*

## 5.11 Eigenvalues and Eigenvectors of Tensors

As we have already seen, decomposing a matrix relies on the spectral properties of the matrix. Similarly, the first indication of a connection between tensor decomposition and tensor eigenvalue problem is contained in (5.7). For example, to approximate a tensor $\mathcal{Y} \in V_1 \otimes \cdots \otimes V_p$ using a rank-1 tensor $\mathcal{X}$, the CP algorithm in (5.7) amounts to solving:

$$\mathbf{Y}_{(n)}(U^{(n+1)} \otimes \cdots \otimes U^{(p)} \otimes U^{(1)} \otimes \cdots \otimes U^{(n-1)}) = \lambda U^{(n)},$$

where $\lambda \in \mathbb{R}$ and $U^{(n)}$ is a unit vector in $\mathbb{R}^{\dim V_n}$. In terms of tensor components, this equation becomes

$$\sum_{i_1,\dots,i_p} Y^{i_1 \dots i_p} (U^{(1)})^{i_1} \cdots (U^{(n-1)})^{i_{n-1}} (U^{(n+1)})^{i_{n+1}} \cdots (U^{(p)})^{i_p} = \lambda (U^{(n)})^{i_n},$$

which is a tensor version of the matrix eigenvalue problem. Similar to symmetric matrices, we can define the eigenvalues and eigenvectors of symmetric tensors using generalized Rayleigh quotients.