Data compression for positive vectors

Amina Benaceur

January 30, 2025

Let N and $\mathcal{N}$ be two integers such that $0 < N \leq \mathcal{N}$. Suppose that we have a family of positive column vectors $(\boldsymbol{a}^1, \dots, \boldsymbol{a}^N) \in \mathbb{R}_+^{\mathcal{N}}$ forming the matrix $\boldsymbol{A} \in \mathbb{R}^{\mathcal{N} \times N}$. We denote X_N the space spanned by positive linear combinations of the vectors $(\boldsymbol{a}^1, \dots, \boldsymbol{a}^N)$, i.e.,

$$X_N = \mathrm{span}_+\{\boldsymbol{a}^1, \dots, \boldsymbol{a}^N\}.$$

Let $n > 0$ be an integer satisfying $n < N$. The goal is to find a set of positive vectors that forms a good approximation space for $(\boldsymbol{a}^1, \dots, \boldsymbol{a}^n)$. Ideally, we search for the set of vectors $(\boldsymbol{b}^1, \dots, \boldsymbol{b}^n)$ satisfying

$$(\boldsymbol{b}^1, \dots, \boldsymbol{b}^n) = \underset{(\boldsymbol{b}^k)_{k=1}^n \in X_N}{\mathrm{argmin}} \ \underset{z \in \mathrm{span}_+\{\boldsymbol{b}^k\}_{k=1}^n}{\mathrm{argmin}} \sum_{i=1}^N \|\boldsymbol{a}^i - z\|^2. \tag{1}$$

In actual practice, the vectors $(\boldsymbol{b}^1, \dots, \boldsymbol{b}^n)$ solution to problem (1) are difficult to compute. We devise an algorithm that builds an alternate set of vectors $(\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^n)$ such that

$$\forall j \in \{1, \dots, N\}, \ \ \exists (\alpha^{j,1}, \dots \alpha^{j,n}) \in \mathbb{R}_+^n : \quad \boldsymbol{a}^j \approx \sum_{k=1}^n \alpha^{j,k} \boldsymbol{\xi}^k. \tag{2}$$

It holds that $\boldsymbol{A}^T \boldsymbol{A} \in \mathbb{R}_+^{N \times N}$. Additionally, if the columns of $\boldsymbol{A}$ are pairwise non-orthogonal, we get $\boldsymbol{A}^T \boldsymbol{A} \in (\mathbb{R}_+^*)^{N \times N}$. The Perron-Frobenius theorem thus ensures the existence of $\lambda^1 \in \mathbb{R}_+^*$ and $\boldsymbol{v}^1 \in (\mathbb{R}_+^*)^{N \times N}$ such that $\boldsymbol{A}^T \boldsymbol{A} \boldsymbol{v}^1 = \lambda^1 \boldsymbol{v}^1$. We then set the first basis vector as $\boldsymbol{\xi}^1 = \boldsymbol{v}^1$. Additionally, the residual vector $\boldsymbol{r}^{j,1}$ associated with $\boldsymbol{a}^j$satisfies

$$\boldsymbol{a}^j = < \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \boldsymbol{\xi}^1 + \boldsymbol{r}^{j,1}. \tag{3}$$

Note that $\boldsymbol{r}^{j,1}$ is not necessarily positive. Hence, we adjust the decomposition of $\boldsymbol{a}^j$ as follows

$$\boldsymbol{a}^j = \left(< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > - \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1} \right) \boldsymbol{\xi}^1 + \boldsymbol{\psi}^{j,1}, \tag{4}$$

Proposition 1. *The decomposition in* (4) *is positive, i.e.*

1. $\boldsymbol{\xi}^1 \in (\mathbb{R}_+^*)^N$;
2. *The coefficient* $< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > - \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1}$ *is positive;*

3. $\boldsymbol{\psi}^{j,1} \in \mathbb{R}_+^N$.

Proof. It holds that $\boldsymbol{\xi}^1 \in (\mathbb{R}_+^*)^N$ by construction.

Second, we prove *2.* By definition, a_i^j is positive for all $i, j \in \{1, \ldots, N\}$. Hence,

$$\forall\, 1 \leq i \leq N: \quad < \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \geq < \boldsymbol{a}^j, \boldsymbol{\xi}^1 > - \frac{a_i^j}{\xi_i^1}.$$

Maximizing over $i \in \{1, \ldots, N\}$, we get

$$< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \geq \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1},$$

whereof the positivity of the coefficient $< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > - \max\limits_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1}$.

Let us now address *3.* It follows from (4) that

$$\begin{aligned}
\forall l \in \{1, \ldots, N\}: \quad \psi_l^{j,1} &= a_l^j - \left(< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > - \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1} \right) \xi_l^1. \\
&= a_l^j - < \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_l^1 + \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1} \xi_l^1. \\
&= - \frac{\left(< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_l^1 - a_l^j \right)}{\xi_l^1} \xi_l^1 + \max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1} \xi_l^1. \\
&= \left(\max_{1 \leq i \leq N} \frac{< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_i^1 - a_i^j}{\xi_i^1} - \frac{\left(< \boldsymbol{a}^j, \boldsymbol{\xi}^1 > \xi_l^1 - a_l^j \right)}{\xi_l^1} \right) \xi_l^1.
\end{aligned}$$

Both factors in the latter equality are positive, whereof the positivity of $\psi_l^{j,1}$ for all $l \in \{1, \ldots, N\}$. It then follows that $\boldsymbol{\psi}^{j,1} \in \mathbb{R}_+^N$, which concludes the proof. □

We can then reiterate the same process at stage k by applying the same strategy to the set of positive vectors $(\boldsymbol{\psi}^{j,1}, \ldots, \boldsymbol{\psi}^{j,N})$. We introduce the matrix $\boldsymbol{\Psi}^{k-1} \in \mathbb{R}_+^{N \times \mathcal{N}}$ whose column vectors are $\boldsymbol{\psi}^k$ for $k \in \{1, \ldots, N\}$. The eigenvector associated with the dominant eigenvalue of the matrix $\boldsymbol{\Psi}^{k-1,T} \boldsymbol{\Psi}^{k-1}$ is denoted $\boldsymbol{\xi}^k$ so that $\boldsymbol{\Psi}^{k-1,T} \boldsymbol{\Psi}^{k-1} \boldsymbol{\xi}^k = \lambda^k \boldsymbol{\xi}^k$. Consequently, at stage $m \in \{0, \ldots, n\}$, such a strategy yields positive decompositions of the form

$$\boldsymbol{a}^j \approx \sum_{k=1}^{m} \alpha^{j,k} \boldsymbol{\xi}^k, \qquad \forall j \in \{1, \ldots, N\}. \tag{5}$$

where, for all $j \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, m\}$,

$$\alpha^{j,k} =< \boldsymbol{\psi}^{j,k-1}, \boldsymbol{\xi}^k > - \max_{1 \leq i \leq N} \frac{< \boldsymbol{\psi}^{j,k-1}, \boldsymbol{\xi}^k > \xi_i^k - \psi_i^{j,k-1}}{\xi_i^k}. \tag{6}$$

We then set the approximation space as $\text{span}_+\{\boldsymbol{A}\boldsymbol{\xi}^1, \ldots, \boldsymbol{A}\boldsymbol{\xi}^n\}$.