

Data compression for positive vectors

AMINA BENACEUR

January 29, 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 $(\mathbf{a}^1, \dots, \mathbf{a}^N) \in \mathbb{R}_+^{\mathcal{N}}$ forming the matrix $\mathbf{A} \in \mathbb{R}^{\mathcal{N} \times N}$. We denote X_N the space spanned by positive linear combinations of the vectors $(\mathbf{a}^1, \dots, \mathbf{a}^N)$, i.e.,

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

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 $(\mathbf{a}^1, \dots, \mathbf{a}^n)$. Ideally, we search for the set of vectors $(\mathbf{b}^1, \dots, \mathbf{b}^n)$ satisfying

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

In actual practice, the vectors $(\mathbf{b}^1, \dots, \mathbf{b}^n)$ 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

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

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

$$\mathbf{a}^j = \langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle \boldsymbol{\xi}^1 + \mathbf{r}^{j,1}. \quad (3)$$

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

$$\mathbf{a}^j = \left(\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle - \max_{1 \leq i \leq N} \frac{\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle \xi_i^1 - a_i^j}{\xi_i^1} \right) \boldsymbol{\xi}^1 + \boldsymbol{\psi}^{j,1}, \quad (4)$$

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

1. $\boldsymbol{\xi}^1 \in (\mathbb{R}_+^*)^N$;
2. The coefficient $\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle - \max_{1 \leq i \leq N} \frac{\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle \xi_i^1 - a_i^j}{\xi_i^1}$ 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 $i, j \in \{1, \dots, N\}$. Hence,

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

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

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

whereof the positivity of the coefficient $\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle - \max_{1 \leq i \leq N} \frac{\langle \mathbf{a}^j, \boldsymbol{\xi}^1 \rangle \xi_i^1 - a_i^j}{\xi_i^1}$.

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

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

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

We can then reiterate the same process at stage k by applying the same strategy to the set of positive vectors $(\boldsymbol{\psi}^{j,1}, \dots, \boldsymbol{\psi}^{j,N})$. We introduce the matrix $\boldsymbol{\Psi}^{k-1} \in \mathbb{R}_+^{N \times N}$ whose column vectors are $\boldsymbol{\psi}^k$ for $k \in \{1, \dots, N\}$. 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 $m \in \{0, \dots, n\}$, such a strategy yields positive decompositions of the form

$$\mathbf{a}^j \approx \sum_{k=1}^m \alpha^{j,k} \boldsymbol{\xi}^k, \quad \forall j \in \{1, \dots, N\}. \quad (5)$$

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

$$\alpha^{j,k} = \langle \boldsymbol{\psi}^{j,k-1}, \boldsymbol{\xi}^k \rangle - \max_{1 \leq i \leq N} \frac{\langle \boldsymbol{\psi}^{j,k-1}, \boldsymbol{\xi}^k \rangle \xi_i^k - \psi_i^{j,k-1}}{\xi_i^k}. \quad (6)$$

We then set define approximation space as $\text{span}_+ \{\mathbf{A}\boldsymbol{\xi}^1, \dots, \mathbf{A}\boldsymbol{\xi}^n\}$ is a