Basic cone geometry and algebra for NMF

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft : December, 1, 2019
Last update : December 9, 2019

Non-negative scaling and non-negative combination

- Given a vector $\mathbf{w} \in \mathbb{R}^m$, the non-negative scaling of $\mathbf{w}$ is a vector $\mathbf{x} \in \mathbb{R}^m$ in the form

$$\mathbf{x} = \alpha \mathbf{w},$$

  where $\alpha \geq 0$.
  Geometric meaning : vector $\mathbf{x}$ is a compressed $(0 < \alpha < 1)$ or extended $(\alpha > 1)$ version of $\mathbf{x}$. If $\alpha = 0$, $\mathbf{w}$ is compressed into a point.

- Given two vector $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^m$, the non-negative combination of $\mathbf{w}_1, \mathbf{w}_2$ is a vector $\mathbf{x} \in \mathbb{R}^m$ in the form

$$\mathbf{x} = \alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{w}_2,$$

  where $\alpha_1 \geq 0, \alpha_2 \geq 0$.
  Geometric meaning : vector $\mathbf{x}$ is within the cone formed by $(\mathbf{w}_1, \mathbf{w}_2)$.

- Other name of non-negative combination : conical combination.

Cone

- Given a set of vectors $W = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$, the cone generated by W is a set that contains all non-negative combination of $\mathbf{w}$ in W. Mathematically

$$\text{cone}(W) = \text{cone}(\mathbf{w}_1, \dots, \mathbf{w}_n) := \left\{ \mathbf{x} \;\middle|\; \mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{w}_i, \alpha_i \geq 0 \, \forall i \right\}.$$

  Let $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^\top$ and $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n]$, we have the equivalent matrix form expression

$$\text{cone}(W) = \text{cone}(\mathbf{W}) := \left\{ \mathbf{x} \;\middle|\; \mathbf{x} = \mathbf{W}\alpha, \alpha \geq 0 \right\},$$

  where usually such α is denoted by $\mathbf{h}$ in NMF literature :

$$\text{cone}(\mathbf{W}) := \left\{ \mathbf{x} \;\middle|\; \mathbf{x} = \mathbf{W}\mathbf{h}, \mathbf{h} \geq 0 \right\}.$$

- The vectors in W that generate the cone are called generators.

Infinite and finite cone

Algebraically

- A cone is called infinite cone if n is infinite.
- A cone is called finite cone if n is finite.
- A finite cone in finite dimensional space always has the matrix form representation. Hence, the equivalent definition of finite cone is that the matrix $\mathbf{W}$ exists (and with size $m < +\infty, n < +\infty$).

Geometrically

- A finite cone is also called polyhedral cone : the base of such cone is a polygon (with finite number of edge).
- An example of infinite cone is the circular cone : the base of such cone is a circle, which is a n-gon with $n = +\infty$.

Extreme ray and simplicial cone

- Given a cone $\mathcal{W} := \text{cone}(\mathbf{W})$, an element in $\mathcal{W}$ is an extreme ray if it cannot be expressed as a non-negative combination of other elements in $\mathcal{W}$. Mathematically, $\mathbf{e} \in \mathcal{W}$ is an extreme ray in $\mathcal{W}$ if

$$\mathbf{e} \notin \Big\{\mathcal{W} - \text{cone}(\mathbf{e})\Big\},$$

  i.e., if you take $\mathbf{e}$ out from $\mathcal{W}$ (and thereby take out $\text{cone}(\mathbf{e})$), $\mathbf{e}$ is no longer in $\mathcal{W}$.
- A cone is simplicial if all the generators of such cone are extreme rays. i.e., all the generator cannot be expressed as non-negative combination of other elements in the cone.
- The order of a simplicial cone is the number of the generator of the cone.
- Important note : zero element is excluded in such definition, as zero element can always be expressed as a non-negative combination of any element in $\mathcal{W}$ (just set the weights to zero). Furthermore, zero element is also not counted as generator.

More on finitely generated simplicial cone

For finitely generated simplicial cone $\mathcal{W}$,

- There exists a matrix $\mathbf{W}$ such that $\mathcal{W} := \text{cone}(\mathbf{W})$
- $\mathbf{W}$ is m-by-n with m, n being finite
- $\mathbf{W}$ has full (column) rank
- The (column) rank of $\mathbf{W}$ $=$ the number of generator of $\mathcal{W}$ $=$ the order of $\mathcal{W}$
- All generator of $\mathcal{W}$ are extreme ray
- All generator of $\mathcal{W}$ cannot be expressed as conical combination of other elements in $\mathcal{W}$

Note. Some statements above are if and only if statements.

Dual cone

Given a cone $\mathcal{W}$, the dual of $\mathcal{W}$ is a set, denoted as $\mathcal{W}^*$, defined as

$$\mathcal{W}* := \left\{ \mathbf{y} \in \mathbb{R}^m \;\middle|\; \langle \mathbf{y}, \mathbf{w} \rangle \geq 0, \forall \mathbf{w} \in \mathcal{W} \right\}$$

Geometrically, $\mathcal{W}$ is the subspace in $\mathbb{R}^m$ formed by the union of the subspaces generated by $\mathbf{y}$ such that the angles between all $\mathbf{y}$ and all $\mathbf{w} \in \mathcal{W}$ are within 90-degree.

Lemmas : Cone inclusion

Given two matrices $\mathbf{X} \in \mathbb{R}^{m \times n_1}$ and $\mathbf{W} \in \mathbb{R}^{m \times n_2}$. We have

$$\text{cone}(\mathbf{X}) \subseteq \text{cone}(\mathbf{W}) \iff \text{all columns of } \mathbf{X} \in \text{cone}(\mathbf{W}).$$

Proof. $(\Leftarrow)$. Assume all columns of $\mathbf{X} \in$ cone$(\mathbf{W})$, then $\alpha_j \mathbf{X}(:, j) \in$ cone$(\mathbf{W})$ for all $\alpha_j \geq 0$ for $j = 1, \dots, n$, which implies[1]

$$\alpha_1 \mathbf{X}(:, 1) + \alpha_2 \mathbf{X}(:, 2) + \cdots + \alpha_n \mathbf{X}(:, n) \in \text{cone}(\mathbf{W}),$$

for all non-negative $(\alpha_1, \dots, \alpha_n)$, meaning cone$(\mathbf{X}) \subseteq$ cone$(\mathbf{W})$.

$(\Rightarrow)$ cone$(\mathbf{X}) \subseteq$ cone$(\mathbf{W})$ means for all non-negative $(\alpha_1, \dots, \alpha_n)$, $\alpha_1 \mathbf{X}(:, 1) + \alpha_2 \mathbf{X}(:, 2) + \cdots + \alpha_n \mathbf{X}(:, n) \in$ cone$(\mathbf{W})$, then by assigning $[\alpha_1, \alpha_2, \dots, \alpha_n]$ as $[1, 0, \dots, 0]$, we have $X(:, 1) \in$ cone$(\mathbf{W})$. Similarly, by assigning $[\alpha_1, \alpha_2, \dots, \alpha_n]$ as $[0, 1, \dots, 0]$, we have $X(:, 2) \in$ cone$(\mathbf{W})$, and so on. Hence all columns of $\mathbf{X} \in$ cone$(\mathbf{W})$. □

(These assignments are valid as cone$(\mathbf{X}) \subseteq$ cone$(\mathbf{W})$ for **all** non-negative $(\alpha_1, \dots, \alpha_n)$)

Note: in this lemma, n_1, n_2 can be different.
i.e., one matrix can be very fat while other one can be very thin.

[1] $\alpha \mathbf{a} \in$ cone$(\mathbf{W})$ and $\beta \mathbf{b} \in$ cone$(\mathbf{W})$ implies $\alpha \mathbf{a} + \beta \mathbf{b} \in$ cone$(\mathcal{W})$, if α, β are non-negative.

Lemma : Dual cone

The dual of $\text{cone}(\mathbf{W})$ for a given matrix $\mathbf{W}$ is

$$\text{cone}^*(\mathbf{W}) = \left\{\mathbf{y} \mid \mathbf{W}^\top \mathbf{y} \geq 0\right\}.$$

Proof.

$$\begin{aligned}
\text{cone}^*(\mathbf{W}) &= \left\{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \geq 0, \mathbf{x} \in \text{cone}(\mathbf{W})\right\} && \text{by definition of dual cone} \\
&= \left\{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \geq 0, \mathbf{x} = \mathbf{W}\mathbf{h}, \mathbf{h} \geq 0\right\} && \text{by definition of cone} \\
&= \left\{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{W}\mathbf{h} \rangle \geq 0, \mathbf{h} \geq 0\right\} && \text{see (*) below} \\
&= \left\{\mathbf{y} \mid \mathbf{h}^\top \mathbf{W}^\top \mathbf{y} \geq 0, \mathbf{h} \geq 0\right\} && \text{see (*) below} \\
&= \left\{\mathbf{y} \mid \mathbf{W}^\top \mathbf{y} \geq 0\right\} \quad \square && \text{see (**) below}
\end{aligned}$$

(*) move $\mathbf{x} = \mathbf{W}\mathbf{h}$ into the dot product
(**) by fact $\mathbf{a} \geq \iff \mathbf{b}^\top \mathbf{a} \geq 0$ for all $\mathbf{b} \geq 0$, and here $\mathbf{b} = \mathbf{h}$, $\mathbf{a} = \mathbf{W}^\top \mathbf{y}$.

Last page - summary

- Definition of cone
- Definition of extreme ray
- Definition of simplicial cone
- Definition of dual cone
- $\text{cone}(\mathbf{X}) \subseteq \text{cone}(\mathbf{W}) \iff \mathbf{X}(:, j) \in \text{cone}(\mathbf{W}) \; \forall j$
- $\text{cone}^*(\mathbf{W}) = \{\mathbf{y} \mid \mathbf{W}^\top \mathbf{y} \geq 0\}$

End of document