IS - NMF

* **Problem** :

We want to solve $\min_{\underline{h}_n \in \mathbb{R}_+^K} C(\underline{h}_n)$ where

$$C(\underline{h}_n) = D_{IS}(\underline{v}_n \mid \underline{\underline{W}}\,\underline{h}_n)$$

$$= \sum_{f=1}^{F} d_{IS}\left([\underline{v}_n]_f \mid [\underline{\underline{W}}\,\underline{h}_n]_f \right)$$

$$= \sum_{f=1}^{F} d_{IS}\left(v_{fn} \mid \sum_{k=1}^{K} w_{fk} h_{kn} \right)$$

$$= \sum_{f=1}^{F} \left[\frac{v_{fn}}{\sum_k w_{fk} h_{kn}} + \ln\left(\sum_k w_{fk} h_{kn} \right) \right] + cst$$

$$= \breve{C}(\underline{h}_n) + \hat{C}(\underline{h}_n) + cst$$

where $\breve{C}(\underline{h}_n) = \sum_{f=1}^{F} \breve{C}_f(\underline{h}_n)$ and $\hat{C}(\underline{h}_n) = \sum_{f=1}^{K} \hat{C}_f(\underline{h}_n)$

with $\breve{C}_f(\underline{h}_n) = \dfrac{v_{fn}}{\sum_k w_{fk} h_{kn}}$ and $\hat{C}_f(\underline{h}_n) = \ln\left(\sum_k w_{fk} h_{kn} \right)$

roposition (auxiliary function to $C(\underline{h}_n)$) :

The function $G(\underline{h}_n | \underline{\tilde{h}}_n)$ defined below is an auxiliary unction to $C(\underline{h}_n)$;

$$G(\underline{h}_n | \underline{\tilde{h}}_n) = \breve{G}(\underline{h}_n | \underline{\tilde{h}}_n) + \hat{G}(\underline{h}_n | \underline{\tilde{h}}_n)$$

with

- $$\breve{G}(\underline{h}_n | \underline{\tilde{h}}_n) = \sum_{f=1}^{F} \left[\sum_{k=1}^{K} \left[\frac{w_{fk} \tilde{h}_{kn}^{2}}{h_{kn} \left(\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n} \right)} \right] v_{fn} \right]$$

- $$\hat{G}(\underline{h}_n | \underline{\tilde{h}}_n) = \sum_{f=1}^{F} \left[\ln\left(\sum_{k=1}^{K} w_{fk} \tilde{h}_{kn} \right) + \frac{\sum_{k=1}^{K} w_{fk} h_{kn} - \sum_{k=1}^{K} w_{fk} \tilde{h}_{kn}}{\sum_{k=1}^{K} w_{fk} \tilde{h}_{kn}} \right]$$

PROOF

Concave part :

→ The condition $\hat{G}(\underline{h}_n, \underline{h}_n) = \hat{C}(\underline{h}_n)$ is trivially met.

→ We will prove that $\hat{C}(\underline{h}_n) \leq G(\underline{h}_n \mid \tilde{\underline{h}}_n)$

by majorizing each term $\hat{C}_f(\underline{h}_n)$.

As the composition of a concave function $(x \mapsto \ln(x))$ and a linear function, $\hat{C}_f(\underline{h}_n)$ is a concave function, so it can be majorized by its tangent (1st order Taylor expansion) at an arbitrary point $\tilde{\underline{h}}_n \in \mathbb{R}_+^K$:

$$\hat{C}_f(\underline{h}_n) \leq \hat{G}_f(\underline{h}_n \mid \tilde{\underline{h}}_n) := \hat{C}_f(\tilde{\underline{h}}_n) + \nabla^T \hat{C}_f(\tilde{\underline{h}}_n)(\underline{h}_n - \tilde{\underline{h}}_n),$$

where ∇^T denotes the transpose of the gradient.

Equality iff $\underline{h}_n = \tilde{\underline{h}}_n$.

From the definition of $\hat{C}_f(\underline{h}_n)$ we can develop :

$$\hat{G}_f(\underline{h}_n \mid \tilde{\underline{h}}_n) = \ln\left(\sum_k w_{fk} \tilde{h}_{kn}\right) + \sum_k \left[\frac{w_{fk}}{\sum_{k'} w_{fk'} \tilde{h}_{k'n}} \left(h_{kn} - \tilde{h}_{kn}\right)\right]$$

$$G_f(\underline{h}_n \mid \tilde{\underline{h}}_n) = \ln\left(\sum_k w_{fk} \tilde{h}_{kn}\right) + \frac{\sum_k w_{fk} h_{kn} - \sum_k w_{fk} \tilde{h}_{kn}}{\sum_k w_{fk} \tilde{h}_{kn}}$$

Finally,

$$\hat{C}(\underline{h}_n) = \sum_f \hat{C}_f(\underline{h}_n) \leq \sum_f \hat{G}_f(\underline{h}_n \mid \tilde{\underline{h}}_n) = \hat{G}(\underline{h}_n \mid \tilde{\underline{h}}_n)$$

which completes the 1st part of the proof.

* Convex part:

→ The condition $\breve{G}(\underline{h}_n \mid \underline{h}_n) = \breve{C}(\underline{h}_n)$ is trivially met.

→ We will show that $\breve{C}(\underline{h}_n) \leq \breve{G}(\underline{h}_n \mid \tilde{\underline{h}}_n)$ by majorizing each term $\breve{C}_f(\underline{h}_n)$.

We introduce the set of variable $\left\{\tilde{\phi}_{k,fn} \in [0,1]\right\}_{k=1}^{K}$ defined by

$$\tilde{\phi}_{k,fn} = \frac{w_{fk} \tilde{h}_{kn}}{\sum_{k'} w_{fk'} \tilde{h}_{k'n}}$$

...h that $\sum_k \tilde{\phi}_{k,fn} = 1$.

Jensen's inequality: if $x \mapsto \varphi(x)$ is convex on $\mathbb{R}_+$, we have $\varphi\left(\sum_k x_k\right) \leq \sum_k \tilde{\phi}_k \varphi\left(\frac{x_k}{\tilde{\phi}_k}\right)$ for $x_1, \dots, x_K \geq 0$ and $\tilde{\phi}_1, \dots, \tilde{\phi}_K \geq 0$ such that $\sum_k \tilde{\phi}_k = 1$.

We have equality iff $\tilde{\phi}_k = \frac{x_k}{\sum_{k'} x_{k'}}$.

As the composition of a convex function $\left(x \mapsto \frac{1}{x} \text{ on } \mathbb{R}_+\right)$ and a linear function, $\breve{C}_f(\underline{h}_n)$ is a convex function.

Using Jensen's inequality we have:

$$\breve{C}_f(\underline{h}_n) \leq \breve{G}_f(\underline{h}_n \mid \underline{\tilde{h}}_n) := v_{fn}\left[\sum_k \tilde{\phi}_{k,fn} \cdot \frac{\tilde{\phi}_{k,fn}}{w_{fk} h_{kn}}\right] = v_{fn}\left[\sum_k \frac{\tilde{\phi}_{k,fn}^2}{w_{fk} h_{kn}}\right]$$

We can inject in this equation the expression of the auxiliary variables $\{\tilde{\phi}_{k,fn}\}_k$ such that:

$$\sigma_f(\underline{h}_n|\tilde{\underline{h}}_n) = \left[\sum_k \frac{\cancel{w_{fk}}\,\tilde{h}_{kn}^2}{\left(\sum_{k'} w_{fk'}\hat{h}_{k'n}\right)^2} \frac{1}{\cancel{w_{fk}}\,h_{kn}}\right] v_{fn}$$

$$= \left[\sum_k \frac{w_{fk}\,\tilde{h}_{kn}^2}{h_{kn}\left(\sum_{k'} w_{fk'}\tilde{h}_{k'n}\right)^2}\right] v_{fn}.$$

Then,

$$\breve{C}(\underline{h}_n) = \sum_f \hat{C}_f(\underline{h}_n) \leq \sum_f \breve{G}_f(\underline{h}_n|\tilde{\underline{h}}_n) = G(\underline{h}_n|\tilde{\underline{h}}_n)$$

which completes the 2^{nd} part of the proof.

Finally,

$$C(\underline{h}_n) = \breve{C}(\underline{h}_n) + \hat{C}(\underline{h}_n) \leq \breve{G}(\underline{h}_n|\tilde{\underline{h}}_n) + \hat{G}(\underline{h}_n|\tilde{\underline{h}}_n) = G(\underline{h}_n|\tilde{\underline{h}}_n)$$

which completes the proof.

$$\frac{\partial G(\underline{h}_n | \tilde{\underline{h}}_n)}{\partial h_{k,n}} = \sum_{f=1}^{F} \left[\frac{w_{fk}}{\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n}} - \frac{v_{fn}}{h_{kn}^2} \times \frac{w_{fk} \tilde{h}_{kn}^2}{\left(\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n} \right)^2} \right] = 0$$

$$\Leftrightarrow h_{kn}^{*} = \left[\frac{\sum_{f=1}^{F} v_{fn} w_{fk} \tilde{h}_{kn}^2 \left(\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n} \right)^{-2}}{\sum_{f=1}^{F} w_{fk} \left(\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n} \right)^{-1}} \right]^{1/2}$$

$$\Leftrightarrow h_{kn}^{*} = \tilde{h}_{kn} \left[\frac{\sum_{f=1}^{F} v_{fn} w_{fk} \left(\sum_{k'=1}^{K} w_{fk'} \tilde{h}_{k'n} \right)^{-2}}{\sum_{f=1}^{F} w_{fk} \left(\sum_{k'=1}^{K} w_{fk'} h_{k'n} \right)^{-1}} \right]^{1/2}$$

Using the fact that $\tilde{h}_{kn}$ is equal to the previous value of h_{kn} according to the MM algorithm, we can summarize the update as follows:

$$h_{kn} \leftarrow h_{kn} \left[\frac{\sum_{f=1}^{F} v_{fn} w_{fk} \left(\underline{\underline{W}}\, \underline{\underline{H}} \right)_{fn}^{-2}}{\sum_{f=1}^{F} w_{fk} \left(\underline{\underline{W}}\, \underline{\underline{H}} \right)_{fn}^{-1}} \right]^{1/2}$$

In matrix form: $\underline{\underline{H}} \leftarrow \underline{\underline{H}} \odot \left[\frac{\underline{\underline{W}}^T \left(\underline{\underline{V}} \odot \left(\underline{\underline{W}}\, \underline{\underline{H}} \right)^{\odot -2} \right)}{\underline{\underline{W}}^T \left(\underline{\underline{W}}\, \underline{\underline{H}} \right)^{\odot -1}} \right]^{\odot 1/2}$