Optimization Problem (1)

$$\min_{\substack{W,H,Q \\ \mathcal{G}\geq 0}} D_\beta(\mathcal{X} \,|\, \mathcal{G}\times_1 W\times_2 H\times_3 Q) + \mu_G \|\mathcal{G}\|_1 + \frac{1}{2}\left(\mu_W\|W\|_F^2 + \mu_H\|H\|_F^2 + \mu_Q\|Q\|\right.$$

where $\mu_G, \mu_W, \mu_H, \mu_Q > 0$

Subproblem in W

Let us first recall: $\mathcal{X}_{(1)} = W \underbrace{\mathcal{G}_{(1)}(H\otimes Q)^T}_{=(\mathcal{G}\times_2 H\times_3 Q)_{(1)}}$

For clarity, let us pose $\begin{cases} V = \mathcal{X}_{(1)} \\ H = (\mathcal{G}\times_2 H\times_3 Q)_{(1)}\end{cases}$. The sub-problem in W becomes:

$$\min_{W\geq 0} D_\beta(V|WH) + \frac{1}{2}\mu_W \|W\|_F^2 \quad (2)$$

It is obvious that (2) is separable w.r.t. rows of W, hence the objective of (2) can be written as follows:

$$\sum_f \left[D_\beta(v_f | w_f H) + \tfrac{1}{2} \mu_W \|w_f\|_F^2\right]$$

We focus on the minimization of one particular row of W, for clarity we remove the concerned subscript "f":

$$\min_{w\geq 0} D_\beta(v|wH) + \tfrac{1}{2}\mu_W\|w\|_F^2 \quad (3)$$

For tackling (3), we use the MM framework (build a convex separable auxiliary function for (3) and minimize it in closed form). The second term is separable. For the first term, we use the auxiliary function from (Févotte et al.) build at the current $\tilde{w}$, denoted $G(w|\tilde{w})$.

Finally we want to solve:

$$\min_{w\geq 0} G(w|\tilde{w}) + \frac{1}{2} \mu_W\|w\|_F^2 \quad (4)$$

To do so: we find the w that cancels the gradient of (4). Since the objective function (4) is separable w.r.t. each entry w_k of w, we focus on solving: $\nabla_{w_k}\left[G(w|\tilde{w}) + \tfrac{1}{2}\mu_W\|w\|_F^2\right] = 0$

For $\beta = 1$: $\iff \sum_m h_{km} - \tilde{w}_k \sum_m h_{km}\frac{v_m}{w_k \tilde{v}_m} + \mu_W w_k = 0$

for $w_k>0$, $\iff \mu_W w_k^2 + \left(\sum_m h_{km}\right) w_k - \tilde{w}_k \sum_m h_{km}\frac{v_m}{\tilde{v}_m} = 0$

that can be solved in closed form. (There is one nonnegative root). Note: $\tilde{v}_m = [\tilde{w} H]_m$

$$\Delta = \text{"} b^2 - 4a.c \text{"}$$
$$= \left(\sum_m h_{km}\right)^2 + 4 \mu_W \tilde{w}_{fk} \sum_m h_{km} \frac{v_m}{\tilde{v}_m}$$

The first root is:

$$w_{fk} = \frac{\sqrt{\left(\sum_m h_{km}\right)^2 + 4 \mu_W \tilde{w}_{fk} \sum_m h_{km} \frac{v_m}{\tilde{v}_m}} - \sum_m h_{km}}{2 \mu_W} \quad (5)$$

which is nonnegative.

Let us pose $C = eH^T$ with e a all-one matrix of size $(x_{(1)})$

$$S = 4 \mu_W \tilde{W} \odot \left(\frac{[V]}{[\tilde{W}H]} H^T \right)$$

$\odot$: hadamard product

(5) can be expressed in matrix form as follows:

$$W = \frac{\left[[C^{.2} + S]^{.1/2} - C \right]}{2 \mu_W} \quad (6)$$

Rmk
- Computing the limit $\mu_W \to 0$ by using Hospital's rule makes (6) tends to original. MU introduced by Lee and Seung in 2000.
- Finding the updates (which are not "multiplicative") corresponds to find the positive root of a polynomial (monovariate) equation. The degree of the polynomial equation is $1 + (2 - \beta)$ for $\beta \in [0, 2]$
  $\Rightarrow$ case by case, no general formula for any β!