

TDNMF 算法损失函数:

$$\min_{W,H} \|Y - WH^T\|_F^2 + \lambda_1(\|W\|_F^2 + \|H\|_F^2) + \lambda_2\|W - T\|_F^2 + \lambda_3 \sum_{i,p=1}^n \|h_i - h_q\|^2 C_{ip} \quad (1)$$

$$\text{s. t. } W \geq 0, H \geq 0$$

推导思路 1: 梯度下降法

首先, 我们考虑无约束条件

$$\min_{W,H} \|Y - WH^T\|_F^2 + \lambda_1(\|W\|_F^2 + \|H\|_F^2) + \lambda_2\|W - T\|_F^2 + \lambda_3 \sum_{i,p=1}^n \|h_i - h_q\|^2 C_{ip} \quad (2)$$

通过“梯度下降法”求迭代更新公式:

$$LF = \|Y - WH^T\|_F^2 + \lambda_1(\|W\|_F^2 + \|H\|_F^2) + \lambda_2\|W - T\|_F^2 + \lambda_3 \sum_{i,p=1}^n \|h_i - h_q\|^2 C_{ip}$$

$$\begin{aligned} &= \text{Tr}[(Y - WH^T)(Y - WH^T)'] + \lambda_1\{\text{Tr}(WW') + \text{Tr}(HH')\} + \\ &\quad \lambda_2\text{Tr}[(W - T)(W - T)'] + \lambda_3\text{Tr}(H^T LH) \\ &= \text{Tr}(YY^T) - 2\text{Tr}(YHW') + \text{Tr}(WH^T HW') + \lambda_1\text{Tr}(WW') + \\ &\quad \lambda_1\text{Tr}(HH') + \lambda_2\text{Tr}(WW') - 2\lambda_2\text{Tr}(WT') + \lambda_2\text{Tr}(TT') + \lambda_3\text{Tr}(H^T LH) \end{aligned} \quad (3)$$

下面开始求梯度:

$$\begin{aligned} \frac{\partial LF}{\partial W} &= -2YH + 2WH'H + 2(\lambda_1 + \lambda_2)W - 2\lambda_2T \\ &= 2((WH'H + (\lambda_1 + \lambda_2)W) - (YH + \lambda_2T)) \end{aligned} \quad (4)$$

$$\frac{\partial LF}{\partial H} = 2((HW'W + \lambda_1H + \lambda_3DH) - (Y^TW + \lambda_3CH)) \quad (5)$$

下面开始使用梯度下降法:

$$\begin{aligned} W_{ik} &\leftarrow W_{ik} - \eta \frac{\partial LF}{\partial W} \\ &\leftarrow W_{ik} + 2\eta \left\{ (YH + \lambda_2T)_{ik} - (WH'H + (\lambda_1 + \lambda_2)W)_{ik} \right\} \end{aligned} \quad (6)$$

公式(6)是无约束情况, 而我们真正需要的是有约束($W \geq 0, H \geq 0$)的情况, 这里我们可以令 $\eta = \frac{W_{ik}}{2(WH'H + (\lambda_1 + \lambda_2)W)_{ik}}$, 便可使其满足约束条件($W \geq 0, H \geq 0$)。因此, 我们可得:

$$W_{ik} \leftarrow \frac{(YH + \lambda_2T)_{ik}}{(WH'H + (\lambda_1 + \lambda_2)W)_{ik}} W_{ik} \quad (7)$$

同理可得:

$$H_{ik} \leftarrow \frac{(Y^TW + \lambda_3CH)_{ik}}{(HW'W + \lambda_1H + \lambda_3DH)_{ik}} H_{ik} \quad (8)$$

推导思路 2：拉格朗日函数法

利用 KKT 互补松弛条件实施非负约束。（1）式的拉格朗日函数为：

	$LF = \ Y - WH'\ _F^2 + \lambda_1(\ W\ _F^2 + \ H\ _F^2) + \lambda_2\ W - T\ _F^2 + \lambda_3 \sum_{i,p=1}^n \ h_i - h_q\ ^2 C_{ip}$ $+ Tr(\Lambda_1 W') + Tr(\Lambda_2 H')$	
	$= Tr[(Y - WH')(Y - WH')'] + \lambda_1\{Tr(WW') + Tr(HH')\} +$ $\lambda_2 Tr[(W - T)(W - T)'] + \lambda_3 Tr(H^T LH) + Tr(\Lambda_1 W') + Tr(\Lambda_2 H')$	
	$= Tr(YY^T) - 2Tr(YHW') + Tr(WH'HW') + \lambda_1 Tr(WW') +$ $\lambda_1 Tr(HH') + \lambda_2 Tr(WW') - 2\lambda_2 Tr(WT') + \lambda_2 Tr(TT') + \lambda_3 Tr(H^T LH) +$ $Tr(\Lambda_1 W') + Tr(\Lambda_2 H')$	(9)

我们有下列的 KKT 条件

	$\Lambda_1 \odot W = 0$	(10)
	$\Lambda_2 \odot H = 0$	(11)

接下来，我们对公式(9)求导可得：

	$\frac{\partial LF}{\partial W} = -2YH + 2WH'H + 2(\lambda_1 + \lambda_2)W - 2\lambda_2 T + \Lambda_1$ $= 2((WH'H + (\lambda_1 + \lambda_2)W) - (YH + \lambda_2 T)) + \Lambda_1$	(12)
	同理可得：	
	$\frac{\partial LF}{\partial H} = 2((HW'W + \lambda_1 H + \lambda_3 DH) - (Y^T W + \lambda_3 CH)) + \Lambda_2$	(13)
	令(12)和(13)为 0，可得：	
	$\Lambda_1 = -2((WH'H + (\lambda_1 + \lambda_2)W) - (YH + \lambda_2 T))$	(14)
	$\Lambda_2 = -2((HW'W + \lambda_1 H + \lambda_3 DH) - (Y^T W + \lambda_3 CH))$	(15)
	公式(14)(15)与(10)(11)相结合得：	
	$((WH'H + (\lambda_1 + \lambda_2)W) - (YH + \lambda_2 T)) \odot W = 0$	(16)
	$((HW'W + \lambda_1 H + \lambda_3 DH) - (Y^T W + \lambda_3 CH)) \odot H = 0$	(17)
	因此，我们可以得到更新公式：	
	$W_{ik} \leftarrow \frac{(YH + \lambda_2 T)_{ik}}{(WH'H + (\lambda_1 + \lambda_2)W)_{ik}} W_{ik}$	(18)
	$H_{ik} \leftarrow \frac{(Y^T W + \lambda_3 CH)_{ik}}{(HW'W + \lambda_1 H + \lambda_3 DH)_{ik}} H_{ik}$	(19)