# Revision: Matrix and Tensor Factorization

Xinyu Chen

February 3, 2022

**2.1** For any partially observed data matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times T}$ with the observed index set $\Omega$, a matrix factorization algorithm can decompose $\boldsymbol{Y}$ into lower dimensional factor matrices $\boldsymbol{W} \in \mathbb{R}^{R \times N}, \boldsymbol{X} \in \mathbb{R}^{R \times T}$, and its loss function can be written as

$$f = \frac{1}{2} \sum_{(i,t) \in \Omega} \left( y_{i,t} - \boldsymbol{w}_i^\top \boldsymbol{x}_t \right)^2 + \frac{\rho}{2} \left( \sum_{i=1}^{N} \|\boldsymbol{w}_i\|_2^2 + \sum_{t=1}^{T} \|\boldsymbol{x}_t\|_2^2 \right), \tag{1}$$

where $\boldsymbol{w}_i \in \mathbb{R}^R$ is the $i$th column of $\boldsymbol{W}$, and $\boldsymbol{x}_t \in \mathbb{R}^R$ is the $t$th column of $\boldsymbol{X}$. The symbol $\|\cdot\|_2$ denotes the $\ell_2$-norm.

1. Obtain the partial derivative with respect to $\boldsymbol{w}_i$, i.e., $\frac{\partial f}{\partial \boldsymbol{w}_i}$.
2. Obtain the partial derivative with respect to $\boldsymbol{x}_t$, i.e., $\frac{\partial f}{\partial \boldsymbol{x}_t}$.
3. How to use Alternating Least Squares (ALS) method to solve the following optimization problem:

$$\min_{\boldsymbol{W}, \boldsymbol{X}} \; \frac{1}{2} \sum_{(i,t) \in \Omega} \left( y_{i,t} - \boldsymbol{w}_i^\top \boldsymbol{x}_t \right)^2 + \frac{\rho}{2} \left( \sum_{i=1}^{N} \|\boldsymbol{w}_i\|_2^2 + \sum_{t=1}^{T} \|\boldsymbol{x}_t\|_2^2 \right). \tag{2}$$

.