

組合せ依存型の Factorization Machines

制約

なぜ目的関数において一部の内積が 0 となるように制約を設定するかについて.

目的関数

データセット D に対して設定される損失関数 J は以下の (1) で表現される.

$$J(D) = \frac{1}{2} \sum_D (y - \hat{y})^2 + \frac{1}{2} C \left\{ \sum_{i=1}^M \sum_{j=i+1}^M (\langle \mathbf{v}_i, \mathbf{v}_j \rangle)^2 + \sum_{i=1}^M \sum_{j=i+1}^M (\langle \mathbf{v}_i, \mathbf{v}_j \rangle)^2 \right\} \quad (1)$$

最適化

$$\frac{\partial}{\partial \theta} = \begin{cases} y - \hat{y} & \text{if } \theta = w_0, \\ (y - \hat{y})x_i & \text{if } \theta = w_i, \\ x_i \sum_{j=1}^{2M+n} v_{j,f} x_j - v_{i,f} x_i^2 + C \left(\sum_{j=i+1}^M \langle \mathbf{v}_i, \mathbf{v}_j \rangle \right) v_{i,f} & \text{if } \theta = v_{i,f} \wedge i \in [1, M], \\ x_i \sum_{j=1}^{2M+n} v_{j,f} x_j - v_{i,f} x_i^2 + C \left(\sum_{j=i+1}^{2M} \langle \mathbf{v}_i, \mathbf{v}_j \rangle \right) v_{i,f} & \text{if } \theta = v_{i,f} \wedge i \in [M+1, 2M], \\ x_i \sum_{j=1}^{2M+n} v_{j,f} x_j - v_{i,f} x_i^2 & \text{otherwise;} \end{cases} \quad (2)$$