An Online Learning Algorithm for Bilinear Models (Supplementary)

Yuanbin Wu YBWU@CS.ECNU.EDU.CN
Shiliang Sun SLSUN@CS.ECNU.EDU.CN
Shanghai Key Laboratory of Multidimensional Information Processing
Department of Computer Science and Technology, East China Normal University

Details of Algorithm 2

Several technical notes on power iteration are about $\|\alpha_t\|, \|\beta_t\|$ and σ_1.

First, when the iteration stops, $\bar{\alpha}_t^{(\tau)}, \bar{\beta}_t^{(\tau)}$ may not have unit length. To normalize them, we also need to compute their norms. Observing that

$$\|\alpha_t\|^2 = \|\alpha_{t-1}\|^2 + 2\langle\alpha_{t-1}, \Delta\alpha_t^{(\tau)}\rangle + \|\Delta\alpha_t^{(\tau)}\|^2,$$

and the last two terms only involve manipulations on sparse vectors, we can maintain the norms efficiently.

Second, it is also possible to avoid explicit normalization $\frac{\alpha_t^{(\tau)}}{\|\alpha_t^{(\tau)}\|}$ by further refining update equations (9)-(12) with $\|\alpha_{t-1}\|, \|\beta_{t-1}\|$ involved.

$$\Delta\alpha^{(\tau)} = \frac{\|\alpha_{t-1}\|}{\sigma_1\|\beta_{t-1}\|}\left(C(\Delta\Phi^t)\beta_{t-1} + \Theta_t\Delta\beta^{(\tau-1)}\right), \quad (1)$$

$$\bar{\alpha}_t^{(\tau)} = \alpha_{t-1} + \Delta\alpha^{(\tau)}, \quad (2)$$

$$\Delta\beta^{(\tau)} = \frac{\|\beta_{t-1}\|}{\sigma_1\|\alpha_{t-1}\|}\left(C(\Delta\Phi^t)^\intercal\alpha_{t-1} + \Theta_t^\intercal\Delta\alpha^{(\tau)}\right), \quad (3)$$

$$\bar{\beta}_t^{(\tau)} = \beta_{t-1} + \Delta\beta^{(\tau)}. \quad (4)$$

When the power iteration stops, we can simply set α_t, β_t as follows (without explicit normalization):

$$\alpha_t = \alpha_{t-1} + \Delta\alpha^{(R)},$$
$$\beta_t = \beta_{t-1} + \Delta\beta^{(R)}.$$

Finally, for σ_1, simple algebra manipulations gives the updates:

$$\sigma_1^t = \frac{1}{\|\alpha_t\|\|\beta_t\|}\left(\sigma_1^{t-1}\|\alpha_{t-1}\|\|\beta_{t-1}\| + C\beta_{t-1}^\intercal\Delta\Phi\alpha_{t-1} + \Delta\beta^\intercal\Theta_t\alpha_{t-1} + \beta_{t-1}^\intercal\Theta_t\Delta\alpha + \Delta\beta^\intercal\Theta_t\Delta\alpha\right).$$