Homework 3 Hints: Mathematics of Deep Learning (EN 580.745)

December 10, 2019

1. Variant of matrix factorization dropout.

(a) First, note that for $\bar{\boldsymbol{u}}\bar{\boldsymbol{v}}^\top \sim \mathcal{D}_{(\boldsymbol{U},\boldsymbol{V})}$, we know by construction that $\mathbb{E}[\Theta(\boldsymbol{U},\boldsymbol{V})\bar{\boldsymbol{u}}\bar{\boldsymbol{v}}^\top] = \boldsymbol{U}\boldsymbol{V}^\top$. Now add and subtract $\boldsymbol{U}\boldsymbol{V}^\top$ in the stochastic objective (3). Expand and simplify using properties of expectation and independence to show

$$\mathop{\mathbb{E}}_{\{\bar{\boldsymbol{u}}_j\bar{\boldsymbol{v}}_j^\top\}_{j=1}^m}\left\|\boldsymbol{X}-\boldsymbol{U}\boldsymbol{V}^\top+\boldsymbol{U}\boldsymbol{V}^\top-\frac{\Theta(\boldsymbol{U},\boldsymbol{V})}{m}\sum_{j=1}^m \bar{\boldsymbol{u}}_j\bar{\boldsymbol{v}}_j^\top\right\|_F^2 = \|\boldsymbol{X}-\boldsymbol{U}\boldsymbol{V}^\top\|_F^2+\frac{1}{m}(\Theta(\boldsymbol{U},\boldsymbol{V})^2-\|\boldsymbol{U}\boldsymbol{V}^\top\|_F^2). \quad (1)$$

Finally, complete the square to show that

$$\|\boldsymbol{X}-\boldsymbol{U}\boldsymbol{V}^\top\|_F^2 - \frac{1}{m}\|\boldsymbol{U}\boldsymbol{V}^\top\|_F^2 = \frac{m-1}{m}\|\tilde{\boldsymbol{X}}-\boldsymbol{U}\boldsymbol{V}^\top\|_F^2 + C, \quad (2)$$

where C is a constant independent of $\boldsymbol{U}$ or $\boldsymbol{V}$.