A Simple Exercise on Matrix Derivatives

Jason D. M. Rennie
jrennie@gmail.com

February 18, 2006

Abstract

This is an exercise to remind myself of how to do simple matrix derivatives.

Thomas Minka has an excellent discussion of derivatives of matrices [1]. Let $U \in \mathbb{R}^{n\times k}$, $V \in \mathbb{R}^{m\times k}$, and $Y \in \mathbb{R}^{n\times m}$. Let

$$J(U,V) = \|UV^T - Y\|^2_{\text{Fro}} + \frac{\lambda}{2}(\|U\|^2_{\text{Fro}} + \|V\|^2_{\text{Fro}}) \tag{1}$$

$$= \sum_{i,j}\left(\sum_a U_{ia}V_{ja} - Y_{ij}\right)^2 + \frac{\lambda}{2}\left(\sum_{i,a} U^2_{ia} + \sum_{j,a} V^2_{ja}\right) \tag{2}$$

Let U_i be the i^{th} row of U, let V_j be the j^{th} row of V, let Y_i be the i^{th} row of Y, and let $V_{\cdot a}$ be the a^{th} column of V. Then,

$$\frac{\partial J}{\partial U_{ia}} = 2\sum_j\left(\sum_a U_{ia}V_{ja} - Y_{ij}\right)V_{ja} + \lambda U_{ia}, \tag{3}$$

$$= 2\sum_j\left(U_iV_j^T - Y_{ij}\right)V_{ja} + \lambda U_{ia}, \tag{4}$$

$$= 2\left(U_iV^T - Y_i\right)V_{\cdot a} + \lambda U_{ia}, \tag{5}$$

$$\frac{\partial J}{\partial U_i} = 2\left(U_iV^T - Y_i\right)V + \lambda U_i, \tag{6}$$

$$\frac{\partial J}{\partial U} = 2\left(UV^T - Y\right)V + \lambda U. \tag{7}$$

Similarly,

$$\frac{\partial J}{\partial V} = 2\left(UV^T - Y\right)^T U + \lambda V. \tag{8}$$

Note that this does *not* follow the convention of swapping indices as discussed in [1].

References

[1] T. Minka. Old and new matrix algebra useful for statistics. http://research.microsoft.com/ minka/papers/matrix/, December 2000.