ML & LinAlg Math Cheat Sheet

October 24, 2017

Contents

1 Notation

Vectors are column vectors denoted by lower-case bolded variables, such that

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}.$$

A row vector is denoted $\boldsymbol{x}^\top = [x_1 \dots x_N]$. A matrix is indicated by a bolded upper-case variable, such that an $N \times M$ matrix is

$$\boldsymbol{A} = \{a_{ij}\} = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_M] = \begin{bmatrix} \boldsymbol{a}_1^\top \\ \vdots \\ \boldsymbol{a}_N^\top \end{bmatrix} = \begin{bmatrix} a_{1,1}^\top & \cdots & a_{1,M} \\ \vdots & \ddots & \vdots \\ a_N^\top & \cdots & a_{N,M} \end{bmatrix}.$$

2 Derivative

2.a Random Vector

$$\nabla_{\boldsymbol{x}}\boldsymbol{y} = [\frac{\partial \boldsymbol{y}}{\partial x_1}, \dots, \frac{\partial \boldsymbol{y}}{\partial x_N}] \tag{1}$$

3 Determinant Operator

3.a Random Properties

For scalar c and $N \times N$ identity matrix I,

$$\det(cI) = c^N.$$

4 Trace Operator

4.a Derivatives

4.a.i $d(\text{tr}(\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A}))/d\boldsymbol{x}$

$$d(\text{tr}(\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A}))/d\boldsymbol{x} = \frac{d}{d\boldsymbol{x}} \sum_i^N \sum_k^N x_i x_k a_{ik}.$$

Consider

$$\frac{d}{dx_j} \sum_i^N \sum_k^N x_i x_k a_{ik} = x_1 a_{j,1} + x_2 a_{j,2} + \cdots \frac{d}{dx_j} \sum_k^N x_j x_k a_{jk} + \cdots x_N a_{j,N}$$

4.b Relation to Determinant