Notes on Information Geometry

Charles Zheng

August 24, 2015

1 Definitions

The α-connection

$$\Gamma^{(\alpha)}_{ij,k}(\theta) = \mathbf{E}_\theta[((\partial_i \partial_j \ell_\theta) - (\partial_i \ell_\theta)(\partial_j \ell_\theta))\partial_k \ell_\theta]$$

2 Exponential families

2.1 Definitions

$$p_\theta(x) = C(x)\exp[\Sigma_i \theta_i t_i(x) - \psi(\theta)]$$

$$\ell_\theta(x) = \log p_\theta(x) = \log C(x) + \sum_i \theta_i t_i(x) - \psi(\theta)$$

$$\psi(\theta) = \log \int_x C(x) e^{\theta^T t(x)} dx$$

$$\eta_i(\theta) = \mathbf{E}_\theta[t_i(X)]$$

$$g_{ij}(\theta) = \mathrm{Cov}_\theta[t_i(X), t_j(X)] = g_{ji}$$

2.2 Properties under natural parameterization

$$\partial_i \triangleq \frac{\partial}{\partial \theta_i}$$

$$\partial_i \psi(\theta) = \eta_i(\theta)$$

$$\partial_i \partial_j \psi(\theta) = g_{ij}(\theta)$$

$$\partial_i \ell(x) = t_i(x) - \eta_i$$

Note that $\partial_i \partial_j \ell(x)$ is constant as a function of x.

$$\partial_i \partial_j \ell(x) = -g_{ij}$$

$$\partial_i \partial_j \partial_k \ell(x) = -\partial_k g_{ij} = -\mathbb{E}(\partial_i \ell)(\partial_j \ell)(\partial_k \ell) \stackrel{D}{=} -T_{ijk}$$

$$T_{ijk} = \partial_k g_{ij} = T_{ikj} = \cdots = T_{kji}$$

2.3 Natural parameterization is e-affine

$$\Gamma^{(1)}_{ij,k}(\theta) = \mathbf{E}_\theta[(\partial_i \partial_j \ell)(\partial_k \ell)] = -g_{ij}\mathbf{E}[\partial_k \ell] = 0$$

2.4 Properties under η parameterization

$$\frac{\partial \eta_i}{\partial \theta_j} = \partial_j(\partial_i \psi) = g_{ij}$$

$$\tilde{\partial}_i \triangleq \frac{\partial}{\partial \eta_i} = \sum_j \frac{\partial \theta_j}{\partial \eta_i} \partial_j = \sum_j \frac{\partial_j}{g_{ij}}$$

$$\tilde{\partial}_i \ell = \sum_k \frac{\partial_k \ell}{g_{ik}}$$

$$\begin{aligned}
\tilde{\partial}_i \tilde{\partial}_j \ell &= \sum_k \tilde{\partial}_j \left(\frac{\partial_k \ell}{g_{ik}} \right) \\
&= \sum_{m,k} \frac{\partial_m}{g_{jm}} \left(\frac{\partial_k \ell}{g_{ik}} \right) \\
&= \sum_{m,k} \frac{1}{g_{jm}} \left(\frac{g_{mk}}{g_{ik}} - \frac{(\partial_k \ell) T_{imk}}{g_{ik}^2} \right)
\end{aligned}$$