A Class of Convex Functions

Jason D. M. Rennie
jrennie@gmail.com

May 13, 2005

Abstract

We show that functions of the form $f(\vec{a}) = \log\left(\sum_{\vec{x}} Q(\vec{x}) e^{\vec{a}^T\vec{x}}\right)$, $Q(\vec{x}) \geq 0\ \forall \vec{x}$, are convex in $\vec{a}$.

Define $P(x) = \frac{Q(x)e^{ax}}{\sum_x Q(x)e^{ax}}$. Let

$$f(a) = \log\left(\sum_x Q(x)e^{ax}\right). \tag{1}$$

Then

$$\frac{\partial f}{\partial a} = \frac{\sum_x xQ(x)e^{ax}}{\sum_x Q(x)e^{ax}} = E_{x\sim P(x)}[x], \tag{2}$$

and

$$\frac{\partial^2 f}{\partial a \partial a} = \frac{\sum_x x^2 Q(x)e^{ax}}{\sum_x Q(x)e^{ax}} - \frac{\left(\sum_x xQ(x)e^{ax}\right)^2}{\left(\sum_x Q(x)e^{ax}\right)^2} = E[x^2] - E[x]^2 = \mathrm{Var}(x). \tag{3}$$

As is well-known, variance is non-negative; hence f is convex in a.

Consider a vector-valued function,

$$f(\vec{a}) = \log \sum_{\vec{x}} Q(\vec{x}) e^{\vec{a}^T\vec{x}}. \tag{4}$$

Define $P(\vec{x}) = \frac{Q(\vec{x})e^{\vec{a}^T\vec{x}}}{\sum_{\vec{x}} Q(\vec{x})e^{\vec{a}^T\vec{x}}}$. Then,

$$\frac{\partial f}{\partial \vec{a}} = \frac{\sum_{\vec{x}} \vec{x} Q(\vec{x}) e^{\vec{a}^T\vec{x}}}{\sum_{\vec{x}} Q(\vec{x}) e^{\vec{a}^T\vec{x}}} = E_{\vec{x}\sim P(\vec{x})}[\vec{x}], \tag{5}$$

and

$$\frac{\partial^2 f}{\partial \vec{a} \partial \vec{a}} = \sum_{\vec{x}} \vec{x}\vec{x}^T P(\vec{x}) - \left(\sum_{\vec{x}} \vec{x} P(\vec{x})\right)\left(\sum_{\vec{x}} \vec{x} P(\vec{x})\right)^T \tag{6}$$

$$= E[\vec{x}\vec{x}^T] - E[\vec{x}]E[\vec{x}]^T = \mathrm{Cov}(\vec{x}). \tag{7}$$

The covariance matrix of any random vector is positive semi-definite, so the vector-valued version of f is convex in $\vec{a}$.