Derivation of Energy Expression Terms

Michael Iuzzolino

1 Energy Functions and Activation Dynamics

An attractor net is characterized by an energy or Lyapunov function (Koiran, 1994):

$$E_i = -\frac{1}{2}x_i W x_i^{\mathrm{T}} - b x_i^{\mathrm{T}} + \sum_j \int_0^{x_{ij}} f^{-1}(\xi) d\xi, \tag{1}$$

where x_i is the network state at some iteration i, j is an index over units in the net, and $f(\cdot)$ is the activation function. With $f \equiv \tanh$, we have:

$$E_i = -\frac{1}{2}x_i W x_i^{\mathrm{T}} - b x_i^{\mathrm{T}} + \sum_j x_{ij} \mathrm{arctanh}(x_{ij}) + \frac{1}{2}\ln(1 - x_{ij}^2) \tag{2}$$

Derivation

The aim is to show the following:

$$\int_0^{x_{ij}} \tanh^{-1}(\xi) d\xi \equiv x_{ij}\mathrm{arctanh}(x_{ij}) + \frac{1}{2}\ln(1 - x_{ij}^2)$$

Beginning with $\int_0^{x_{ij}} \tanh^{-1}(\xi) d\xi$, we integrate by parts, where $\int u dv = uv - \int v du$.

Let $u = \tanh^{-1}(\xi)$ and $dv = d\xi$. Then,

$$\begin{aligned} u &= \tanh^{-1}(\xi) & v &= \xi \\ du &= \frac{1}{1-\xi^2} d\xi & dv &= d\xi \end{aligned}$$

Following $\int u dv = uv - \int v du$,

$$\xi \tanh^{-1}(\xi)\Big|_0^x - \int_0^x \frac{\xi}{1-\xi^2} d\xi \tag{3}$$

Now we make a substitution: $u = 1 - \xi^2 \Rightarrow \xi = \sqrt{1-u}$. Then $du = -2\xi d\xi \Rightarrow d\xi = \frac{du}{-2\sqrt{1-u}}$.

Evaluating the first term of Eq. 3 and making the above substitution into the second term of Eq. 3 (along with algebraic simplifications and readjustment of integration bounds), we achieve:

$$x\tanh^{-1}(x) + \frac{1}{2}\int_1^{1-x^2} \frac{1}{u} du$$

Since $\int du/u = ln(u) + C$,

$$x\tanh^{-1}(x) + \frac{1}{2}\ln(u)\Big|_1^{1-x^2} = x\tanh^{-1}(x) + \frac{1}{2}\left[\ln(1-x^2) - \ln(1)\right]$$

Finally,

$$\int_0^{x_{ij}} \tanh^{-1}(\xi) d\xi = x\tanh^{-1}(x) + \frac{1}{2}\ln(1-x^2) \quad \square$$