

Week 9 Reading

Si Kai Lee sl3950

November 14, 2016

I read this article last semester and didn't really understand it, but this time I found it much easier to read. I finally understood the reverse referred to in the second last line of the last paragraph of Section 2.3: it meant that placing little mass in $q(\cdot)$ in areas where $p(\cdot)$ has more mass is penalised less. This also helped to reinforce the notion that $KL(q||p)$ leads to mode-matching while $KL(p||q)$ moment-matching.

Section 4 really hit home the importance of choosing distributions that are in the exponential family as such distributions makes the calculation of updates much easier and helped me understand in how Dave showed conjugacy in exponential families as I missed that bit of the class. I would like to better understand how the natural gradient of the ELBO is derived as well as the reasons for the conditions needed to guarantee the convergence of stochastic optimisation algorithms and explore the possibility of including natural gradient in black-box VI.