Probabilistic Graphical Models 10-708
Homework 5 (Extra Credit): Due April 30, 2014 at 4 pm

Directions. Note that this homework is only two problems, both of which are extra credit. Hence, if you do not do this homework, your grade will not be penalized.

This homework assignment covers the material presented in Lectures 19-24. To submit your assignment, please upload a pdf file containing your writeup to Canvas by 4 pm on Wednesday, April 30th. We highly encourage that you type your homework using the LaTeX template provided on the course website, but you may also write it by hand and then scan it.

Important Note. You homework writeup should be saved as a pdf file. Before submitting your assignment, please double check that everything you want to include shows up in the compiled pdf. You must also make sure that this document is legible. This means that if you don't use LaTeX, your handwriting must be neat and your answers must be organized. Next, place all of your code in a directory and compress it into a zip file. Do not place tex files or anything else inside of this directory. Finally, when submitting the assignment, please separately attach both the pdf file and the zip file to your Canvas submission. This makes it much easier for us to grade the non-programming questions because your writeup will load on the page in the grading tool that we use. If you do not follow these instructions, e.g. if you place your writeup inside of the zip file or write illegibly, we will take 5 points off of your final homework grade, and possibly more if we cannot read your answers to certain questions.

1 Posterior of the Dirichlet Process [10 points]

1. (10 points) Let H be a distribution over Θ and let α be a positive scalar. For any finite, measurable partition $A_1, \ldots, A_r$ of Θ, G is defined to be a Dirichlet process with base distribution H and concentration parameter α, denoted by $G \sim \mathrm{DP}(\alpha, H)$, if

$$(G(A_1), \ldots, G(A_r)) \sim \mathrm{Dir}(\alpha H(A_1), \ldots, \alpha H(A_r)). \tag{1}$$

   Suppose we have observations $X_1, \ldots, X_n$, which we assume are drawn from G. Assuming we have the prior $G \sim \mathrm{DP}(\alpha, H)$, derive the posterior distribution for $G|X_1, \ldots, X_n$.

2 Hilbert Space Embeddings [10 points]

We discussed in class that Hilbert Space Embeddings are attractive because certain probability "rules" also hold for the analogous RKHS operators. In class we discussed the RKHS version of the sum rule. It is highly recommended you fully understand the materials in Lecture 23 before doing this question.

Let A and B be random variables. In this question you will prove that $\boldsymbol{\mathcal{C}}_{BA} = \boldsymbol{\mathcal{C}}_{B|A}\boldsymbol{\mathcal{C}}_{AA}$. This is the RKHS analog of $\mathbb{P}[B, A] = \mathbb{P}[B|A]\mathbb{P}[A]$.

We will assume that both random variables A, B are embedded in RKHS $\mathcal{F}$. The corresponding feature functions for $\mathcal{F}$ will be indexed by ϕ. Thus, just to clarify notations/definitions:

$$
\begin{aligned}
\boldsymbol{\mathcal{C}}_{BA} &= \mathbb{E}_{BA}[\phi_B \otimes \phi_A] \qquad \text{RKHS analog of } \boldsymbol{\mathcal{P}}[B, A] \\
\boldsymbol{\mathcal{C}}_{AA} &= \mathbb{E}_{AA}[\phi_A \otimes \phi_A] \qquad \text{RKHS analog of } \boldsymbol{\mathcal{P}}[A, A]
\end{aligned}
$$

1. Assuming A, B are binary variables, write the rule $\mathbb{P}[B, A] = \mathbb{P}[B|A]\mathbb{P}[A]$ in matrix form.
2. Prove that $\boldsymbol{\mathcal{P}}[B, A] = \boldsymbol{\mathcal{P}}[B|A]\boldsymbol{\mathcal{P}}[\oslash A]$ in general using expectations and δ indicator vectors (as done on Lecture 23 Slide 17 for the sum rule).
3. Now prove the RKHS version: $\mathcal{C}_{BA} = \mathcal{C}_{B|A}\mathcal{C}_{AA}$.