Machine Learning Midterm

This TWO-SIDED exam is open book. You may bring in your homework, class notes and textbooks to help you. You will have 1 hour and 15 minutes. Write all answers in your blue book. Please make sure YOUR NAME is on your exam. Square brackets [] denote the points for a question or sub-question. ANSWER ALL FOUR QUESTIONS FOR FULL CREDIT

1. **Probability** [25] Two variables x and y are startistically independent if $p(x, y) = p(x)p(y)$.

   The variance of a variable x is defined by:

   $$\mathsf{var}(x) = \int\int (x - E[x])^2 dx$$

   Use the the fact that

   $(x+y-E[x+y])^2 = (x-E[x])^2+(y-E[y])^2+2(x-E[x])(y-E[y])$

   to show that

   $$\mathsf{var}(x + y) = \mathsf{var}(x) + \mathsf{var}(y)$$

2. **Entropy**

   (a) [5]Given that the elements of a source $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ are independent, how would we express $p(\mathbf{s})$ in terms of its components?

   (b) [15] Given that the **Negative entropy** of $p(\mathbf{s})$ is defined as:

   $$E[\log p(\mathbf{s})]$$

   Show how the expression for maximizing $l(W)$ used in ICA can be interpreted as negative entropy when using $p(\mathbf{s})$ instead of g'.

   Remember that

   $$l(W) = \sum_{i=1}^{n} \log g'(Wx) + \log |W|$$

3. **Support vector machines**

   (a) [5] Given three support vectors $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_3$,such that

   $$\mathbf{w}^T\mathbf{x}_1 + b = 1$$

   $$\mathbf{w}^T\mathbf{x}_2 + b = -1$$

   $$\mathbf{w}^T\mathbf{x}_3 + b = -1.$$

   (b) [10]*Draw a diagram* illustrating how their constraint can be visualized geometrically.

   (c) [15] Using the method of Lagrange multipliers, derive an expression for the offset b.

4. **Sampling**

   (a) [5] *BRIEFLY:* Why is MCMC better than rejection sampling?

   (b) [5] Name one constraint that ,if satisfied, would guarantee that Gibbs sampling could be used.

   (c) [15] In Gibbs sampling it is sometimes better to use over-relaxtion where instead of using the sample z_i that has mean μ_i and variance σ_i, one uses $\hat{z}_i$ defined by

   $$\hat{z}_i = \mu_i + \alpha(z_i - \mu_i) + \sigma_i(1 - \alpha^2)^{1/2}\nu$$

   where ν is a gaussian random variable with zero mean and unit variance and α is a parameter $-1 < \alpha < 1$. Show that the variance of $\hat{z}_i = \sigma$.

   HINT: use the result of Problem 1.