Background Knowledge Quiz

1. Your name:

2. What year and program are you in?

3. Are you taking this course for credit, auditing, or on the waiting list?

Gaussians

4. If $p(x) = \mathcal{N}(x|\mu, \sigma^2)$,

    (a) For some $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$, can $p(x) < 0$?

    (b) For some $x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$, can $p(x) > 1$?

5. If $p(x) = \mathcal{N}(x|\mu, \Sigma)$ with $x \in \mathbb{R}^D, \mu \in \mathbb{R}^D, \Sigma \in \mathbb{R}^{D \times D}$, (a multivariate Gaussian),

    (a) What is the computational complexity (the asymptotic time cost) of evaluating $p(x)$?

    (b) What restrictions are there on Σ in order for it to be a valid covariance matrix?

Derivatives

6. If A is a fixed matrix and $\mathbf{x}$ is a vector, what is $\frac{\partial (A\mathbf{x})_i}{\partial \mathbf{x}_j}$?

7. Given a composition of functions $f(x) = a(b(c(x)))$, we can evaluate its derivative using the chain rule - just multiply together the Jacobian of each function. What is the fastest order to multiply this product of Jacobians $J_a \times J_b \times J_c$, if $f(x)$ is a vector-input, scalar-output function?

8. How could one form an unbiased estimate of $\nabla_x \int f(x, \theta) p(\theta) d\theta$ using samples from $p(\theta)$, and derivatives of f?

Distributions

9. In the natural exponential family of distributions, $p(x|\theta) = f(x)g(\theta) \exp\{x\theta\}$. What must $g(\theta)$ be in order for $p(x|\theta)$ to be a valid probability distribution?

10. One way to specify a Categorical (discrete) distribution using an unconstrained vector $\mathbf{x} \in \mathbb{R}^D$ is with the softmax function: $p(y = c|\mathbf{x}) = \frac{\exp\{x_c\}}{\sum_{c'=1}^{D} \exp\{x_{c'}\}}$:

    (a) What could go wrong numerically in evaluating $p(y = c|\mathbf{x})$ if some elements of $\mathbf{x}$ are large?

    (b) How could one fix this?