Homework 4: 10-36/702 due April 19

1. **Simulation**

   Let
$$X_1, \ldots, X_n \sim \frac{1}{2}N(\mu_1, 1) + \frac{1}{2}N(\mu_2, 1).$$

   (a) Consider the prior
$$\pi(\mu_1, \mu_2) \propto 1.$$

   Show that the posterior in improper (i.e. the posterior does not have finite integral). Hint: we can write the model as:
$$Z_i \sim \text{Bernoulli}(1/2)$$
$$X_i|Z_i \sim Z_i Y_i + (1 - Z_i) W_i$$

   where $Y_i \sim N(\mu_1, 1)$ and $W_i \sim N(\mu_2, 1)$. Write
$$\pi(\mu_1, \mu_2 | X_1, \ldots, X_n) = \sum_z \pi(\mu_1, \mu_2 | X_1, \ldots, X_n, z)\pi(z|X_1, \ldots, X_n)$$

   where $z \in \{0, 1\}^n$. Consider what happens when $z = (0, \ldots, 0)$.

   (b) Now use the prior
$$\pi(\mu_1, \mu_2) = \phi(\mu_1; 0, \tau)\phi(\mu_2; 0, \tau)$$

   where ϕ is the Gaussian distribution and $\tau > 0$ is a fixed constant. Write a Gibbs sampling algorithm to find the posterior for (μ_1, μ_2).

   (c) Generate data $n = 100$ observations with $\mu_1 = -3$ and $\mu = 3$. Take $\tau = 10$. Use your Gibbs sampling algorithm to find the posteriors $\pi(\mu_1|X_1, \ldots, X_n)$ and $\pi(\mu_2|X_1, \ldots, X_n)$.

   (d) Now use a random walk MCMC algorithm to find the posterior. Compare your result to part (c).

2. **Subgradients**

   Derive the subdifferential of the ℓ_∞ norm, i.e. $\partial \| \cdot \|_\infty(\mathbf{x})$ where $\mathbf{x}$ is a d-dimensional vector.

3. **Max of compressed Gaussians**

   In deriving the lasso sparsistency result, we needed a bound on the max of compressed projections of noise. Assume that the noise $\epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ for $i = 1, \ldots, k$ and derive a bound on $\|\Phi^T \epsilon\|_\infty$ for a deterministic projection matrix $\mathbf{\Phi} \in \mathbb{R}^{k \times d}$ that satisfies the RIP condition.

4. **Fenchel dual of the Lasso problem**

   Given a primal optimization problem

$$\inf_{\mathbf{x}} \{f(\mathbf{x}) + g(\mathbf{A}\mathbf{x})\}$$

   its fenchel dual optimization problem is given as

$$\sup_{\mathbf{z}} \{-f^*(\mathbf{A^T z}) - g^*(-\mathbf{z})\}.$$

   (a) Derive the fenchel dual problem of the primal lasso optimization problem:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{\Phi x}\|^2 + \lambda \|\mathbf{x}\|_1$$

   (b) Re-write the dual problem as a constrained optimization problem of the form:

$$\max_{\mathbf{z} \in A} h(\mathbf{z})$$

   where A corresponds to a ball in some norm and $h(\mathbf{z})$ is a quadratic in $\mathbf{z}$.

   (c) Argue why the dual problem can be solved by projecting $\mathbf{y}$ onto a polyhedron. You don't need to actually solve the dual problem.

   Notice: normally $dim(\mathbf{x}) \gg dim(\mathbf{y})$, so $\Phi^T \Phi$ may not be inverted.