1. **Community Detection Using MAP**

It may be helpful to work on this problem in conjunction with the relevant lab. The *stochastic block model (SBM)* defines the random graph $G(n, p, q)$ consisting of two communities of size $\frac{n}{2}$ each, such that the probability an edge exists between two nodes of the same community is $p$, and the probability an edge exists between two nodes in different communities is $q < p$. The goal of the problem is to exactly determine the two communities, given only the graph.

Show that the MAP estimate of the two communities is equivalent to finding the *min-bisection* or *balanced min-cut* of the graph, the split of $G$ into two groups of size $\frac{n}{2}$ that has the minimum edge weight across the partition. Assume that any assignment of the communities is a priori equally likely.
2. Poisson Process MAP

Customers arrive to a store according to a Poisson process with rate 1. The store manager learns of a rumor that one of the employees is sending every other customer to the rival store, so that every odd-numbered customer $1, 3, 5, \ldots$ is sent to the rival store deterministically.

Let $X = 1$ be the hypothesis that the rumor is true and $X = 0$ the rumor is false, assuming that both hypotheses are equally likely. Suppose a customer arrives to the store at time 0. After that, the manager observes $T_1, \ldots, T_n$, where $T_i$ is the time of the $i$th subsequent sale, $i = 1, \ldots, n$. Derive the MAP rule to determine whether the rumor was true or not.
3. **BSC: MLE and MAP**

You are testing a digital link that corresponds to a BSC with some error probability $\varepsilon \in [0, 0.5]$.

a. Suppose that you observe an input bit $X$ and an output bit $Y$. Calculate the MLE of $\varepsilon$.

b. You are now told that the inputs $X_1, \ldots, X_n$ are i.i.d. bits distributed as Bernoulli(0.6). Suppose that you observe $n$ outputs $Y_1, \ldots, Y_n$. Calculate the MLE of $\varepsilon$. 

4. Exponential MLE, MAP, and Hypothesis Testing

Let $X$ be Exponentially distributed with rate 1. Given $X$, the random variable $Y$ is Exponentially distributed with parameter $X$.

a. Find MLE($X \mid Y$).

b. Find MAP($X \mid Y$).

c. For given $c > 1$, maximize $\mathbb{P}(\hat{X} = 1 \mid X = 1)$ subject to $\mathbb{P}(\hat{X} = 1 \mid X = c) \leq 5\%.$
5. **Bayesian Estimation of Exponential Distribution**

We have seen the MLE (non-Bayesian perspective) and MAP estimation (Bayesian perspective). In this problem, we will introduce the fully Bayesian approach to statistical estimation.

Suppose that $X$ is Exponential with unknown rate $\Lambda$. As a Bayesian practitioner, you have a prior belief that the random variable $\Lambda$ is equally likely to be $\lambda_1$ or $\lambda_2$.

Now, you collect one sample $X_1$ from $X$.

a. Find the posterior distribution $P(\Lambda = \lambda_1 \mid X_1 = x_1)$.

b. If we were using the MLE or MAP rule, we would choose a single value $\lambda$ for $\Lambda$, sometimes called a *point estimate*. This amounts to saying $X$ has Exponential distribution with rate $\lambda$. In the Bayesian approach, we will instead keep the full information of the posterior distribution of $\Lambda$, and we compute the distribution of $X$ as

$$f_X(x) = \sum_{\lambda \in \{\lambda_1, \lambda_2\}} f_{X \mid \Lambda}(x \mid \lambda) \cdot P(\Lambda = \lambda \mid X_1 = x_1).$$

Note that we do not necessarily have an Exponential distribution for $X$ anymore. Compute $f_X(x)$ in closed form.

c. From the previous part, you may have guessed that the fully Bayesian approach is often computationally intractable, which is one of the main reasons why point estimates are common in practice. Supposing that $\lambda_1 > \lambda_2$, compute the MAP estimate for $\Lambda$, and calculate $f_X(x)$ again using the MAP rule.
6. **Laplace Prior and $\ell^1$-Regularization**

Suppose you draw $n$ i.i.d. data points $(x_1, y_1), \ldots, (x_n, y_n)$, where the true relationship is given by $Y = WX + \varepsilon$ for $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. In other words, $Y$ has a linear dependence on $X$ with additive Gaussian noise. Further suppose that $W$ has a *Laplace* prior distribution

$$f_W(w) = \frac{1}{2\beta} e^{-|w|/\beta}.$$

Show that finding the MAP estimate of $W$ given the data points $\{(x_i, y_i)\}_{i=1}^n$ is equivalent to minimizing the cost function

$$J(w) = \sum_{i=1}^n (y_i - wx_i)^2 + \lambda |w|.$$  

(You should determine what $\lambda$ is.) This is interpreted as a one-dimensional $\ell^1$-regularized least-squares criterion, also known as LASSO.