ECE C143A/C243A, Spring 2018 **Final**

Department of Electrical and Computer Engineering Prof. J.C. Kao
University of California, Los Angeles TAs: T. Monsoor, X. Jiang, X. Yang

UCLA True Bruin academic integrity principles apply.
Open: Four pages of notes.
Closed: Book, computer, internet.
3:00-6:00pm.
Thursday, 14 Jun 2018.

State your assumptions and reasoning.
No credit without reasoning.
Show all work on these pages.

Name: ____________________

Signature: ____________________

ID#: _________________

Problem 1	______	/	25
Problem 2	______	/	30
Problem 3	______	/	45
Problem 4	______	/	30
Problem 5	______	/	40
Problem 6	______	/	30
BONUS 7	______	/	10 bonus points
Total	______	/	200 points + 10 bonus

1. (25 points) Basic neuroscience.

   (a) (10 points) True / False. State whether each statement is true or false (1 point each). If a statement is false, you need to briefly correct it in the space below the statement to receive full credit (uncorrected false statements receive 0.5 points).

      i. Consider designing a prosthesis that decodes plan activity (i.e., as in HW #4, you pick one of K targets based on plan neural activity during a delay period). In general, it would be better to record from the dorsal premotor cortex (PMd) than primary motor cortex (M1) for this application.

      ii. During action potential generation, voltage-gated potassium ion (K^+) channels generally open before voltage-gated sodium ion (Na^+) channels.

      iii. Even after paralysis or ALS, motor cortex still encodes signals related to intended movements.

      iv. Nodes of Ranvier are capable of regenerating action potentials because these nodes are especially dense in voltage-gated sodium ion channels.

      v. At rest, there is more sodium outside a neuron than inside. If sodium ion channels suddenly opened, the sodium diffusion force (chemical equilibrium driving force) AND the sodium drift force (electric field driving force) would drive sodium ions inside the cell.

      vi. The sodium-potassium ion pump moves ions against their concentration gradients.

      vii. Consider a new drug that reduces membrane resistance. This drug would make action potential generation faster.

viii. If voltage-gated sodium ion channels did not enter a refractory period, then the action potential amplitude could exceed the equilibrium (Nernst) potential of sodium, E_{Na^+}.

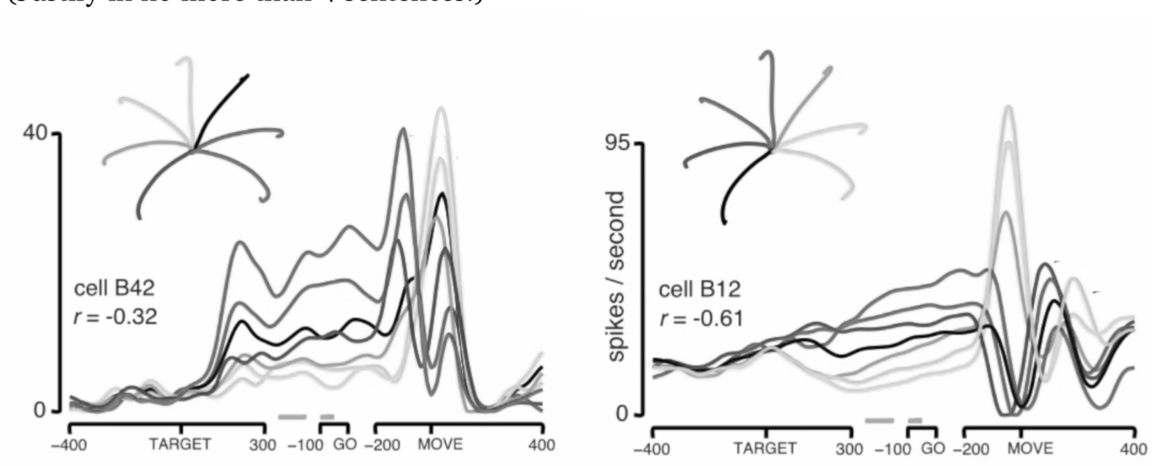
ix. An alien neuron with time-varying firing rate but no refractory period, modeled perfectly by an inhomogeneous Poisson process, will NOT have exponentially distributed interarrival (ISI) times.

x. If two neurons have exponential interarrival times, and thus memoryless ISIs, then neuron 1 and neuron 2's spike times are independent.

(b) (5 points) Wikipedia states: "Multiple sclerosis (MS) is a demyelinating disease in which the insulating covers of nerve cells in the brain and spinal cord are damaged." Explain why demyelination of neurons in the brain would lead to neurological deficits. In particular, how does demyelination affect the propagation of action potentials? (Justify in no more than 5 sentences.)

(c) (5 points) Consider depolarizing a neuron. In class, we saw voltage-clamp experiments that showed when a sub-threshold voltage (which does not trigger an action potential) was applied, the input resistance of the neuron was linear and, in steady-state, was time invariant. However, when an above-threshold voltage (triggering an action potential) was applied, the input resistance of the neuron was nonlinear and time varying. Explain why the input resistance of the neuron becomes nonlinear and time varying when a large enough voltage is applied. (Justify in no more than 4 sentences.)

(d) (5 points) Below, we show two neuron PSTHs during a delayed reach task. The x-axis is time in the trial. The y-axis is the neuron's firing rate. The PSTHs are gray-scale colored by reach direction (upper left panels correspond to the 7 unique reach trajectories for each color). Which, if any, of these two neurons change their preferred direction through time? (Justify in no more than 4 sentences.)



2. (30 points) Poisson processes. Suppose that the number of spikes arriving at an electrode can be modeled by a homogeneous Poisson process with rate 40 spikes per hour. Of these spikes, 10% of the spikes have high amplitude and 90% of the spikes have low amplitude. Suppose that the types of spikes are independent.

   (a) (5 points) Find the probability that in one hour, at least one high amplitude spike is recorded by the electrode.

(b) (5 points) Suppose that during the first hour of observation, exactly 10 high amplitude spikes have been recorded by the electrode. Find the expected number of low amplitude spikes that have been recorded by the electrode during the same period.

(c) (10 points) Find the probability that during the first hour of observation, 4 high amplitude spikes are recorded during the first 10 minutes and 2 low amplitude and 1 high amplitude spikes are recorded during the last 15 minutes of the observation.

(d) (10 points) Suppose that in the first hour of observation, exactly 50 spikes have been recorded by the electrode. Find the probability that among these spikes, there were exactly 5 high amplitude spikes and 45 low amplitude spikes.

3. (45 points) Gaussian Naive Bayes.

A probabilistic generative model for classification comprises class-conditional densities $P(\mathbf{y}|\mathcal{C}_k)$ and class priors $P(\mathcal{C}_k)$, where $\mathbf{y} \in \mathbb{R}^D$ denotes the firing rates of D neurons and $k = 1, 2, \ldots, K$ denotes the class of the reach the monkey was performing (e.g., left, right, up, down, etc.). In homework 4, we derived the maximum likelihood parameter estimates for the Gaussian class-specific covariance model

$$\mathbf{y}|\mathcal{C}_k \sim \mathcal{N}(\mu_k, \Sigma_k)$$

In this problem, we will derive the maximum likelihood parameter estimates for the Gaussian Naive Bayes model, which assumes that:

$$\mathbf{y}|\mathcal{C}_k \sim \mathcal{N}(\mu_k, \Sigma_k), \;\; y_1|\mathcal{C}_k \perp\!\!\!\perp y_2|\mathcal{C}_k \cdots \perp\!\!\!\perp y_D|\mathcal{C}_k$$

where y_i is the ith element of the vector $\mathbf{y}$ for $i = 1, 2, \cdots, D$. This is called a Gaussian Naive Bayes model, since the y_i's are independent Gaussian random variables given $\mathcal{C}_k$.

To keep notation consistent, we will use the following conventions:

- Consider the class of trial i (denoted t_i). We denote the probability that t_i is in class j by π_j, i.e., $\pi_j = P(t_i = j)$.
- We denote the firing rate of neuron m for class j as y_{mj}.
- We denote the standard deviation of neuron m for class j as σ_{mj}.
- Recall that for a scalar random variable, $x \sim \mathcal{N}(\mu, \sigma^2)$, its probability density function is:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$$

(a) (5 points) In the Gaussian Naive Bayes model, the covariance matrix Σ_k is diagonal. Explain why this is the case given our model assumptions.

(b) (10 points) Assume we have N pairs of independent training samples

$$(\mathbf{y}_1, t_1), (\mathbf{y}_2, t_2), \cdots, (\mathbf{y}_N, t_N)$$

where $\mathbf{y}_i \in \mathbb{R}^D$ and $t_i \in \{1, 2, \cdots, K\}$. Write the log-likelihood of observing the training data under the Gaussian Naive Bayes model. You do not have to expand out Gaussian distributions.

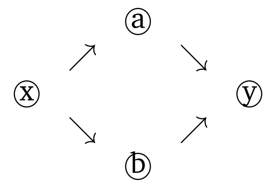
(c) (5 points) Based on your answer to part (b), write the parameter set θ. In the parameter set, you may only list scalars.

(d) (5 points) Write the ML expression for π_j. You do not have to show calculations; however, if you do not show work, you must explain how you arrived at your solution.

(e) (7 points) Derive the ML expression for μ_{mj}. Even if you know intuitively what the answer should be, you need to show your derivation.

(f) (13 points) Derive the ML expression for σ_{mj}. Even if you know intuitively what the answer should be, you need to show your derivation.

4. (30 points) Graphical models. Consider the following graph:



We will determine if x and y are conditionally independent given a and b.

(a) (5 points) Intuitively, are x and y conditionally independent given a and b? Justify your intuition with an argument in words (not math). You may use prior results we derived in class.

(b) (5 points) What do we need to show mathematically to determine if x and y are conditionally independent given a and b? Your answer should be an equation.

(c) (10 points) Use the law of total probability and the chain rule to show that, for this graph:

$$p(x|a,b) = \frac{p(x)p(a|x)p(b|x)}{p(a,b)}$$

(d) (10 points) Determine if $x \perp\!\!\!\perp y | a, b$.

Hint: You may use the result from part c, even if you were unable to show it. If you are familiar with concepts like d-separation or the Bayes ball, you may not use them to show this. Instead, you should show this by simplifying probabilities.

5. (40 points) Optimal linear estimator (OLE) versus inverse optimal linear estimator (IOLE).

Assume we have kinematic data $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K]$ and neural data $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K]$. In the OLE, we solve for the matrix $\mathbf{L}$ such that:

$$\mathbf{x}_k = \mathbf{L}\mathbf{y}_k$$

In the IOLE, we solve for the matrix $\mathbf{B}$ such that:

$$\mathbf{y}_k = \mathbf{B}\mathbf{x}_k$$

(a) (5 points) Write the solution for $\mathbf{L}$ in terms of $\mathbf{X}$ and $\mathbf{Y}$ that minimizes the sum of squared errors:

$$\sum_{i=1}^{K} \|\mathbf{x}_k - \mathbf{L}\mathbf{y}_k\|^2$$

You do not need to show any work.

(b) (5 points) Write the solution for $\mathbf{B}$ in terms of $\mathbf{X}$ and $\mathbf{Y}$ that minimizes the sum of squared errors:

$$\sum_{i=1}^{K} \|\mathbf{y}_k - \mathbf{B}\mathbf{x}_k\|^2$$

You do not need to show any work.

(c) (5 points) Assume we are decoding 2D velocity from 20 neurons, so that $\mathbf{X} \in \mathbb{R}^{2\times K}$ and $\mathbf{Y} \in \mathbb{R}^{20\times K}$. Write the dimensionality of the $\mathbf{L}$ and $\mathbf{B}$ matrix.

(d) (5 points) Your colleague accidentally zeros out the firing rates of neuron 10. Hence, the 10th row of $\mathbf{Y}$ is now all zeros. Will he run into an error in `Python` when he tries to solve for $\mathbf{L}$? Why? (Justify in no more than 4 sentences.)

(Assume your colleague does not use the `np.linalg.pinv` command, but only uses the `np.linalg.inv` command. i.e., if he wants to calculate `np.linalg.pinv(Y)` he will instead compute $\mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1}$ by writing it out explicitly using the `np.linalg.inv` command, i.e., `Y.T.dot(np.linalg.inv(Y.dot(Y.T)))`.

(e) (5 points) Answer part (c), but now for the $\mathbf{B}$ matrix. That is, will your colleague run into a problem when he tries to solve for $\mathbf{B}$? Why? (Justify in no more than 4 sentences.)

(f) (5 points) Your colleague decides to remove neuron 10 entirely, so now $\mathbf{Y} \in \mathbb{R}^{19\times K}$. He wants to compare his solution for $\mathbf{L}$ and $\mathbf{B}$ to your solution with the following key difference: you did not zero out neuron 10's firing rates. Consider the first column of $\mathbf{L}$, which maps the data from neuron 1 to $\mathbf{x}_k$. Will your first columns be the same? Why? (Justify in no more than 4 sentences.)

(g) (5 points) Now consider the first row of $\mathbf{B}$, which maps $\mathbf{x}_k$ to neuron 1's firing rate. Will your first rows be the same? Why? (Justify in no more than 4 sentences.)

(h) (5 points) Your colleague is tired of OLEs and decides to instead work with Kalman filters. He's trying to understand the meaning of $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$. He asks: in a system where $\mathbf{x}_k$ is 1D, is $\Sigma_{k|k}$ larger than $\Sigma_{k|k-1}$? How do you answer your colleague? While you may justify your answer with math, you should provide intuition as to why the answer you arrive at makes sense.

6. (30 points) Principal components analysis. Plotted below are the firing rates of Neuron 1 (X_1) and Neuron 2 (X_2) for an unlabeled reaching task.

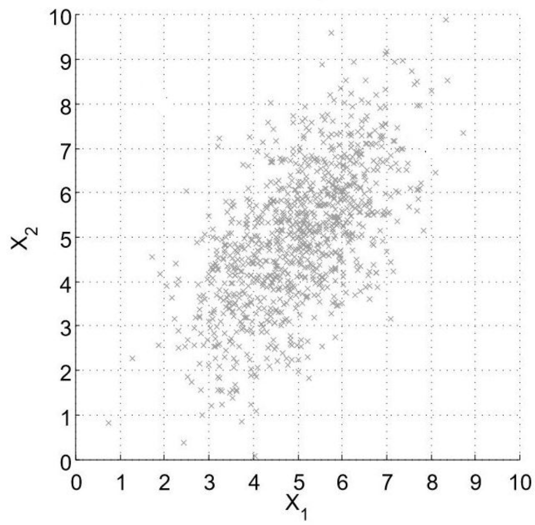


Figure 1: Firing rates of neurons for an unlabeled reaching task

(a) (5 points) Estimate (eye ball) the mean firing rate of Neuron 1 ($\mathbb{E}[X_1]$) and Neuron 2 ($\mathbb{E}[X_2]$) from Figure 1. Round to the nearest integer.

(b) (5 points) Are the firing rates of Neuron 1 and Neuron 2 positively correlated, negatively correlated, or uncorrelated? Please justify your answer in at most two sentences.

(c) (5 points) We define v_1 and v_2 as the directions of the first and second principal component, with $||v_1|| = ||v_2|| = 1$. Sketch and label v_1 and v_2 on Figure 2. In your sketch, the vectors should originate from the point $\begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \end{bmatrix}$.

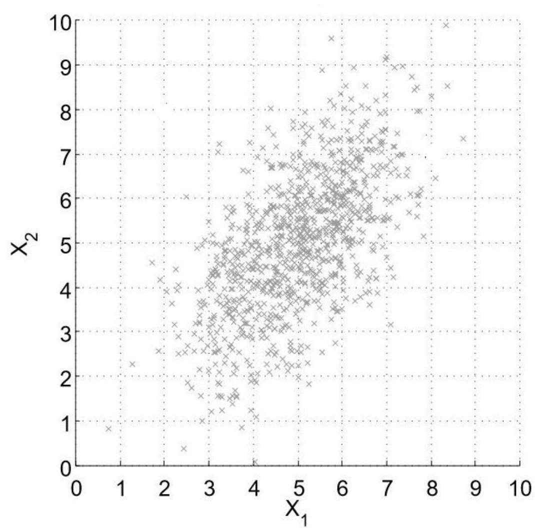


Figure 2: Firing rates of neurons for an unlabeled reaching task

(d) (5 points) Let's consider the points A and B shown in Figure 3

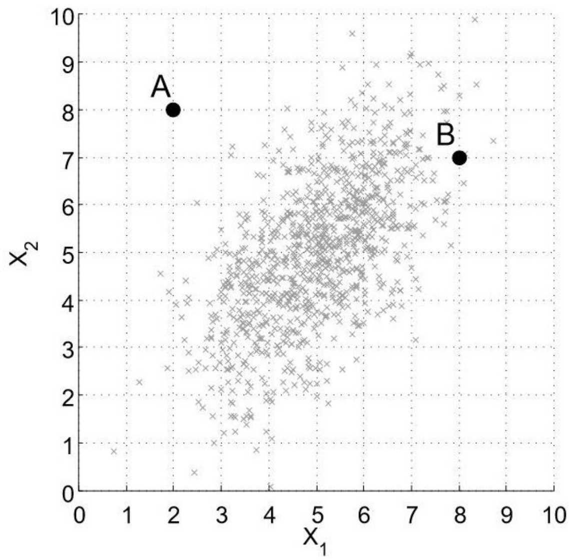


Figure 3: Firing rates of neurons for an unlabeled reaching task

Which point (A or B) would have higher reconstruction error after projecting onto the first principal component v_1? Justify your answer graphically.

(e) (10 points) Let's define the random variables Z_1 and Z_2 in the following manner

$$Z_1 = (\mathbf{X} - \mathbb{E}[\mathbf{X}])^T v_1$$
$$Z_2 = (\mathbf{X} - \mathbb{E}[\mathbf{X}])^T v_2$$

where $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$. Express $\mathbf{cov}(Z_1, Z_2)$ as a function of v_1, v_2 and Σ_X, where Σ_X is the covariance matrix of the random vector $\mathbf{X}$. What do Z_1 and Z_2 represent?

7. (10 points) Bonus question.
A neuroscientist can do neural recordings from two locations, A and B. If he records from location A, then the probability of recording a high amplitude spike is p and a low amplitude spike is $1-p$. If he records from location B, then the probability of recording a high amplitude spike is q and a low amplitude spike is $1-q$. Also, the probability of recording from location A and B are π and $1-\pi$ respectively.

The scientist picks a location for the recording and the sequence of spikes recorded are given below:

$$X = \{1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1\}$$

where $X_i = 1$ denotes a high amplitude spike and $X_i = 0$ denotes a low amplitude spike. For this problem, you may assume that the spike types are conditionally independent.

(a) (2 points) Assuming that the location of the recording is the latent variable, write the parameter set θ.

(b) (6 points) Write the E step for inferring the location of the recording.

(c) (2 points) Answer if the following statement is true or false: iterating between the E-step and M-step will always converge to a local optimum of θ (which may or may not also be a global optimum). Explain your answer in 1-2 sentences.