Number:

Problem	Points	Score
1 (a)	15	
1 (b)	10	
1 (c)	15	
2 (a)	15	
2 (b)	15	
2 (c)	10	
2 (d)	10	
3 (a)	10	
Total	100	

Notes:

1. The exam is closed books and notes. You are allowed one 8 1/2” x 11” double-sided sheet of notes.

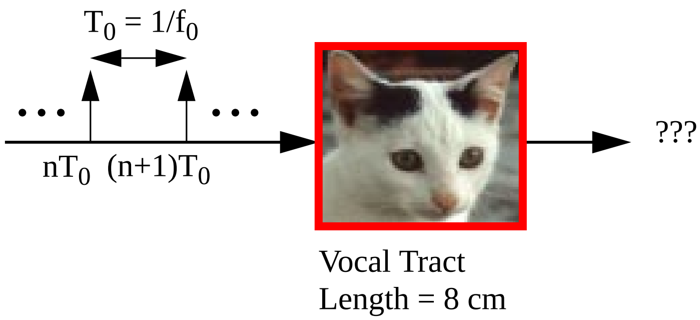
2. Please indicate clearly your answer to the problem by some form of highlighting (underlining).

3. Your solutions must be legible and easy to follow. If I can’t read it or understand it, it is wrong. Random scribbling will not receive credit.

4. Please show ALL work. Answers with no supporting explanations or work will be given no credit.

5. Several problems on this exam are fairly open-ended. Since the evaluation of your answers is obviously a subjective process, we will use a market place strategy in determining the grade. Papers will be rank-ordered in terms of the quality of the solutions, and grades distributed accordingly.

1. Consider the system shown to the right — let's call it a "cat synthesizer". Assume the properties of the cat's vocal tract are the same as a human vocal tract, but the length is shorter. Assume a fundamental frequency (f_0) of 100 Hz. Assume the input is a periodic pulse train as shown.



(a) Design a system to separate the excitation signal from the vocal tract frequency response. Be as specific as possible (e.g., numbers, frequency responses, etc.).

(b) Explain the inaccuracies of this model if a comparable system were used to process real human speech.

(c) Suppose you observe the following features vectors from this system:

$$x[n] = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Design a prewhitening transformation that decorrelates the data and represents an orthonormal transformation. Demonstrate that this transformation works by processing the above data.

2. A system ("black box") outputs the sequence "HTT". Consider these two models:

Assume the probability of starting in each state (initial probabilities) are equal.

(a) Apply the principle of dynamic programming to find which model was the most likely model to generate this sequence.

(b) Next, considering all possible state sequences that could have produced this data, find the model that was most likely to have generated this data.

(c) Assume that "H" represents a value of "0", and "T" represents a value of "1". Also assume the system is capable of outputting continuous-valued outputs (all values in the range [0,1]). Use linear prediction to find the best model of the form $y(n) = ay(n-1) + w(n)$.

(d) Analyze the differences between these three approaches. Do not simply list their features. Discuss what types of assumptions these models make, and why one might be more appropriate than the other.

3. Consider a nonlinear model of a feature vector sequence: $y = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c$,

where $\bar{x}$ is the input vector, $\bar{a}$, $\bar{b}$, and c are parameters, and y is an output variable that indicates which of two classes $\bar{x}$ belongs to. Assume y is a probability in the range [0,1] and represents the probability $\bar{x}$ belongs to the first of two classes.

(a) Explain how you might estimate the optimal values of these parameters using a statistical approach covered in the class lectures. Defend the merits of your approach.