Review Questions, Exam 1

1. What was our definition of learning? (You can paraphrase it) How did it differ from the dictionary's definition?

2. Give a definition of superstitious behavior. Does this fit with our definition of learning?

3. Estes' probability matching experiment involved using a signal light followed by 1 of 2 other lights. The goal was to predict which of those would turn on. What was the outcome of the experiment? How is this related to the card experiment we did in class?

4. (Referring to the previous problem) Why might this be a desirable way to predict the outcomes? Does it maximize our reward (reward being a correct prediction)?

5. What was the $N-$armed bandit problem?

6. Describe (in words) the greedy algorithm and the $\epsilon-$ greedy algorithm. Which is probably a better strategy?

7. Describe in words the softmax strategy. Be sure to include appropriate formulas, and describe what the parameter τ does.

   **NOTE THE ERROR, p. 37: The limits being computed (under the second item after (2.) should be as $\tau \to 0$, not as $\tau \to \infty$.

8. What was the pursuit strategy for the $N-$armed bandit? How does that relate to a "Win-Stay, Lose-Shift" strategy of learning?

9. What was Hebb's postulate for learning? (You can paraphrase it).

10. What was the overall effect of a linear neural network (i.e., what function was being modeled)?

11. There were three update rules we discussed for the weight matrix. What were they? (You only need to state the update for the weight matrix, not the bias update).

12. Matlab Questions:

    (a) What's the difference between a script file and a function?

    (b) What does the following code fragment produce?

    ```
    Q=[1 3 2 1 3];
    idx=find(Q==max(Q));
    ```

    (c) What is the difference between `x=rand;` and `x=randn;`

    (d) What will P be:

```
x=[0.3, 0.1, 0.2, 0.4];
P=cumsum(x);
```

(e) What is the Matlab code that will:

i. Plot $x^2 - 3x$ using 500 points, for $x \in [-1, 4]$

ii. Same plot, but use a red curve.

iii. Compute the variance of data in a vector $\boldsymbol{x}$ (possibly varying in length). You can't use `var`!

iv. Compute the covariance of data in a vector $\boldsymbol{x}$, and $\boldsymbol{y}$ of the same, but possibly varying length. You can't use `cov`!

13. We said that the Fourier transform breaks up a complex waveform into a sum of simple waves. How did it do that? Why does this work? (Give a general description in the continuous case).

14. In the continuous setting, what is the full Fourier expansion of a function in $C[0, 2\pi]$? How are the coefficients computed?

15. If $f(x) = x, g(x) = \mathrm{e}^x$, then give the inner product of f and g (assume these are in $C[0, 2\pi]$).

16. If $f(x) = x^2$ on $[0, 2\pi]$, compute the Fourier coefficients a_0, a_3 and b_3.

17. Suppose we have a function defined on $[0, 2\pi]$, and it has been sampled at 64 points (starting at 0, as usual). I input this data into Matlab's `fft` command, and this is what I get out: All zeros, except for:

$$F(1) = 32, F(3) = -64i, F(6) = 96, F(60) = 96, F(63) = -64i$$

What was the original function?

18. Assume $f(x)$ is real. What is the smallest period that should be used in the discrete Fourier transform?

19. What are the discrete Fourier basis vectors, if $N = 4$?

20. Take the inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$\boldsymbol{x} = [3 - i, 2 + i, -i]^T, \boldsymbol{y} = [1 - i, 1 - i, 3]^T$$

21. What does the power spectrum plot?

22. Compute the mean and variance: $1, 2, 9, 6$

23. Compute the covariance between the data sets:

-2	-1	0	1	2
-6	-3	0	3	6

24. What is the definition of the correlation coefficient? Geometrically, what is the interpretation?

25. What is the definition of the covariance matrix to X? What does the (i, j)th term of the covariance matrix represent?

26. If $f(x, y) = 2xy + x^2 - 4y$, in which direction does f change most rapidly at $(-1, 1)$?

27. If $f(3, 1) = [-1, 3]^T$, and
$$Df(3,1) = \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}$$
give a linear approximation to $f(2, 2)$.

28. Linearize the function
$$\boldsymbol{f}(t) = \begin{bmatrix} \cos(\pi t) \\ t^2 + 2t \end{bmatrix}$$
at the point $t = 1$.

29. Let $f(x, y) = 3x^2 + xy - y^2 + 3x - 5y$
    (a) At $(1, 1)$, in which direction is f increasing the fastest?
    (b) Compute the Hessian of f.
    (c) Rewrite f in the form $\frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$ for appropriate $A, \boldsymbol{b}, c$. Is the choice of A unique?
    (d) Find the stationary point of f.
    (e) Is there a max or min at the stationary point? Hint: Consider the original version of f.

30. Find the orthogonal projection of the vector $\boldsymbol{x} = [1, 0, 2]^T$ to the plane defined by:
$$G = \left\{ \alpha_1 \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} \text{ such that } \alpha_1, \alpha_2 \in \mathbb{R} \right\}$$
Determine the distance from $\boldsymbol{x}$ to the plane G.

31. If $[\boldsymbol{x}]_{\mathcal{B}} = (3, -1)^T$, and $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \right\}$, what was $\boldsymbol{x}$ (in the standard basis)?

32. If $\boldsymbol{x} = (3, -1)^T$, and $\mathcal{B} = \left\{ \begin{bmatrix} 6 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$, what is $[\boldsymbol{x}]_{\mathcal{B}}$?

33. Let $\boldsymbol{a} = [1, 3]^T$. Find a square matrix A so that $A\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{x}$ onto the span of $\boldsymbol{a}$.

34.