REVIEW QUESTIONS, CH 1-5, Modeling

You may prepare and bring a 3×5 card containing notes and/or formulas, and calculators are allowed. The exam will typically be in two parts, one in class and one part that you may take home and work on.

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. What was the $N-$armed bandit problem? In particular, what were the two competing goals, and why were they "competing"?
4. In the $N-$armed bandit problem, how were the estimates of the payoffs, $Q_t(a)$, calculated?
5. There were four "strategies" that we implemented as algorithms to solve the $N-$armed bandit problem. What were they? Be sure to give formulas where appropriate.
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.
8. What was the pursuit strategy (or "Win-Stay, Lose-Shift") for the $N-$armed bandit? Again, include appropriate formulas and describe what β does.
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. What was the final update rule we used for the linear neural network? You may define the weight matrix as W.
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. Compute the variance of data in a vector $\boldsymbol{x}$ (possibly varying in length). You can't use `var`!

iii. 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 had a homework problem where we classified some points in the plane into 4 classes. We built a 2×2 matrix W and a vector $\boldsymbol{b}$ so that so that:

$$W\boldsymbol{x} + \boldsymbol{b} = \begin{bmatrix} \pm 1 \\ \pm 1 \end{bmatrix}$$

depending on whether we had Class 1, 2, 3 or 4. We then plotted those points where $W\boldsymbol{x} + \boldsymbol{b} = 0$. Why did we do that? Geometrically, what did this set of points look like?

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

15. Compute the covariance between the data sets:

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

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

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

18. 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.

19. 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)?

20. 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 $[\boldsymbol{x}]_{\mathcal{B}}$?

21. 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 $\boldsymbol{a}$.

22. Determine the projection matrix that takes a vector $\mathbf{x}$ and outputs the projection of $\mathbf{x}$ onto the plane whose normal vector is $[1,1,1]^T$.

23. Find (by hand) the eigenvectors and eigenvalues of the matrix A:

$$A = \begin{bmatrix} 5 & -1 \\ 3 & 1 \end{bmatrix}, \qquad A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

24. (Referring to the previous exercise) We could've predicted that the eigenvalues of the second matrix would be real, and that the eigenvectors would be orthogonal. Why?

25. Compute the evals and evecs of $A^T A$ if

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}$$

26. Compute the orthogonal projector to the span of $\mathbf{x}$, if $\mathbf{x} = [1,1,1]^T$.

27. Show that $\dfrac{\mathbf{x}\mathbf{x}^T}{\|\mathbf{x}\|^2}$ is a projector, and is an orthogonal projector.

28. Let

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

Find $[\mathbf{x}]_U$. Find the projection of $\mathbf{x}$ into the subspace spanned by the columns of U. Find the distance between $\mathbf{x}$ and the subspace spanned by the columns of U.

29. Show that $\mathrm{Null}(A) \perp \mathrm{Row}(A)$.

30. Show that, if X is invertible, then $X^{-1}AX$ and A have the same eigenvalues.

31. How do we "double-center" a matrix of data?

32. True or False, and give a short reason:

    (a) An orthogonal matrix is an orthogonal projector.
    (b) If the rank of A is 3, the dimension of the row space is 3.
    (c) If the correlation coefficient between two sets of data is 1, then the data sets are the same.

(d) If the correlation coefficient between two sets of data is 0, then there is no functional relationship between the two sets of data.

(e) If U is a 4×2 matrix, then $U^T U = I$.

(f) If U is a 4×2 matrix, then $UU^T = I$.

(g) If A is not invertible, then $\lambda = 0$ is an eigenvalue of A.

(h) Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 0 \end{bmatrix}$$

Then the rank of AA^T is 2.

33. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be the eigenvectors of $A^T A$, where A is $m \times n$.

(a) True or false? The eigenvectors form an orthogonal basis of $\mathbb{R}^n$.

(b) Show that, if $\mathbf{x} \in \mathbb{R}^n$, then the i^{th} coordinate of $\mathbf{x}$ is $\mathbf{x}^T \mathbf{v}_i$.

(c) Let $\alpha_1, \ldots, \alpha_n$ be the coordinates of $\mathbf{x}$ with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
Show that

$$\|\mathbf{x}\|_2 = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2$$

(d) Show that $A\mathbf{v}_i \perp A\mathbf{v}_j$

(e) Show that $A\mathbf{v}_i$ is an eigenvector of AA^T.