# Review Questions, Exam 1 (Math 350, Spr 23)

**Addition to the Review:** I'll allow you to have one half page of notes (8.5" x 5.5") written on one side.

1. What was the $N-$armed bandit problem? In particular, what were the two competing goals, and why were they "competing"?
2. In the $N-$armed bandit problem, how were the estimates of the payoffs, $Q_t(a)$, calculated?
3. 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.
4. Suppose $Q = [-0.5, 0, 0.5, 1.0]$. Use the softmax selection technique with $\tau = 0.1$ to compute the probabilities of choosing each machine.
5. If $Q_1 < Q_2 < Q_3 < Q_4$ for 4 machines, how do the probabilities change (under softmax) as $\tau \to 0$? As $\tau \to \infty$?
6. Suppose we play with three machines, and machine 3 is chosen and gives a big payout (enough to make $Q_t(3)$ the maximum). Update the probabilities for win-stay, lose-shift, if they are: $P_1 = 0.3, P_2 = 0.5, P_3 = 0.2$ and $\beta = 0.3$.
7. 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. Compute the variance of data in a vector $\boldsymbol{x}$ (possibly varying in length). You can't use `var`!
      ii. 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`!
8. What is the definition of the covariance matrix to $X$ (say that $X$ has $p$ vectors in $\mathbb{R}^n$, and $X$ is $n \times p$). You can define it by saying what the $(i, j)$th term of the covariance matrix represents.
9. 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 $G$.
10. 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)?
11. 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 $[\boldsymbol{x}]_{\mathcal{B}}$?

12. 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 $\boldsymbol{a}$.

13. Determine the projection matrix $P$ in each case:

    (a) that projects $\mathbf{x}$ onto vector $\mathbf{a}$ (included for completeness, you answered this in the previous question):

    (b) that projects $\mathbf{x}$ onto the column space of a general matrix $A$ that has full rank:

    (c) that projects $\mathbf{x}$ onto the column space of a matrix $U$, where the columns of $U$ are orthonormal.

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

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

15. (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?

16. Compute the SVD of the matrix $A$ below.

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

17. Show that $\text{Null}(A) \perp \text{Row}(A)$.

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

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

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

    (a) If the rank of $A$ is 3, the dimension of the row space is 3.

    (b) If the correlation coefficient between two sets of data is 1, then the data sets are the same.

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

    (d) If $U$ is a $4 \times 2$ matrix, then $U^T U = I$.

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

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

    (g) Let

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

    Then the rank of $AA^T$ is 2.

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

    (a) Show that if $\lambda_i$ is a non-zero eigenvalue of $A^T A$, then it is also a non-zero eigenvalue of $AA^T$.

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

    (c) Show that, if $\mathbf{x} \in \mathbb{R}^n$, then the $i^{\text{th}}$ coordinate of $\mathbf{x}$ (with respect to the eigenvector basis) is $\mathbf{x}^T \mathbf{v}_i$.

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

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

I'll allow you to show it just using just two vectors, $\mathbf{v}_1, \mathbf{v}_2$.

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

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

22. Given data:

| $x$ | $-1$ | 0 | 1 |
|---|---|---|---|
| $y$ | 2 | 1 | 1 |

(a) Give the matrix equation for the *line of best fit.*

(b) Compute the normal equations.

(c) Solve the normal equations for the slope and intercept.

23. Use the data in Exercise (22) to find the parabola of best fit: $y = ax^2 + bx + c$. (NOTE: Will you only get a least squares solution, or an actual solution to the appropriate matrix equation?)

24. Suppose $\mathbf{x}$ is a vector containing $n$ real numbers, and we understand that $m\mathbf{x} + b$ is Matlab-style notation (so we can add a vector to a scalar, done component-wise).

(a) Find the mean of $\mathbf{y} = m\mathbf{x} + b$ in terms of the mean of $\mathbf{x}$.

(b) Show that, for fixed constants $a, b$, $\text{Cov}(\mathbf{x} + a, \mathbf{y} + b) = \text{Cov}(\mathbf{x}, \mathbf{y})$

(c) If $\mathbf{y} = m\mathbf{x} + b$, then find the covariance and correlation coefficient between $\mathbf{x}$ and $\mathbf{y}$.

25. Given the SVD of a matrix $A$, how do we compute the pseudoinverse of the matrix $A$? Be as specific as you can.

26. Consider the underdetermined "system of equations": $x + 3y + 4z = 1$. In matrix-vector form $A\mathbf{x} = \mathbf{b}$, write the matrix $A$ first.

(a) What is the dimension of each of the four fundamental subspaces?

(b) Find bases for the four fundamental subspaces.

27. (SVD) Given that the SVD of a matrix was given in Matlab as:

```
>> [U,S,V]=svd(A)
U =
   -0.4346   -0.3010    0.7745    0.3326   -0.1000
   -0.1933   -0.3934    0.1103   -0.8886   -0.0777
    0.5484    0.5071    0.6045   -0.2605   -0.0944
    0.6715   -0.6841    0.0061    0.1770   -0.2231
    0.1488   -0.1720    0.1502   -0.0217    0.9619
S =
    5.72         0          0
       0      2.89          0
       0         0          0
       0         0          0
       0         0          0
V =
    0.2321   -0.9483    0.2166
   -0.2770    0.1490    0.9493
    0.9324    0.2803    0.2281
```

(a) Which columns form a basis for the null space of $A$? For the column space of $A$? For the row space of $A$?

(b) How do we "normalize" the singular values? In this case, what are they (numerically)?

(c) What is the rank of $A$?

(d) How would you compute the pseudo-inverse of $A$ (do not actually do it):

(e) Let $B$ be formed using the first two columns of $U$. Would the matrix $B^T B$ have any special meaning? Would $BB^T$?