Math 51 Homework 2

Due Friday July 1, 2016 by 1 pm

Instructions: Complete the following problems. Late homework will not be accepted. Please be sure to review the expectations for your submitted homework outlined online (such as: always including your name and ID number on the homework, stapling your homework, and guidelines for write-ups which will receive full credit). *Make sure to submit your homework to the correct person; (if you are in Section 01, submit to Zev, and if in Section 02, submit to Valentin).*

Part I: Book problems: From Levandosky's *Linear Algebra*, do the following exercises:

- Section 4: #17, 19
- Section 5: #2, 15
- Section 6: #2, 6
- Section 7: #2
- Section 8: #1, 4, 13, 18, 24
- Section 9: #3*, 4abc, 5, 12 (*For #3, your answer should be one or more linear relations involving the entries of **b**).
- Section 10: #11, 12, 17, 21

Part II: Non-book problems:

1. There are 7000 undergraduate students at Stanford. Let M be the 7000×7000 matrix whose ij entry is 1 if student i and student j are Facebook friends and 0 if they are not. (Here we assume that Facebook does not permit a person to be a friend of themselves, so all the diagonal entries of M are zero.) Let $\mathbf{u}$ be the vector in $\mathbb{R}^{7000}$ each of whose entries is 1. What does the vector $M\mathbf{u}$ represent?

2. Suppose you are enrolled in Math 101, a 30-student, project-based course that is graded with a 10% participation component, two short papers worth 25% each, and a final presentation worth 40%. At the end of the quarter, your instructor makes a matrix G with 30 rows and 4 columns. Row i of G contains first the participation score, then the two paper scores, and then the final presentation score for student i (all out of 100%).

(a) Let $\mathbf{v}$ be the vector in $\mathbb{R}^4$ whose entries are 0.1, 0.25, 0.25, and 0.4. What does the vector $G\mathbf{v}$ represent?

(b) Let $\mathbf{w}$ be the vector in $\mathbb{R}^{30}$ each of whose entries is 1/30. What does the dot product of $G\mathbf{v}$ and $\mathbf{w}$ represent?

3. You work for a popular movie rental company, *Hulficks*, and are designing a recommender system to predict what movies a user will like (i.e., rate highly), based on his/her past ratings.

Your company has n different genres of movies, numbered 1 through n. Any particular movie can fall into multiple genres. To each movie we can define its *genre vector*, say $\mathbf{m} \in \mathbb{R}^n$ as follows: if the movie falls into genre number i then the ith component of $\mathbf{m}$ is 1; otherwise it is 0.

Your group has determined that users seem to have internal ratings (also known as *preferences*) of different movie genres, and that (on average), users will rate a given movie m by adding up their internal preferences for the genres that m belongs to. For the purpose of your recommender system, you assume that all users behave this way (and not just on average).

Given a particular user, we can assign them a *preference vector*, say $\mathbf{u} \in \mathbb{R}^n$, whose ith component is the user's internal rating for genre number i.

(a) Given a user with preference vector $\mathbf{u}$, and a particular movie with genre vector $\mathbf{m}$, express the user's rating of this movie in terms of vector operations.

(b) Suppose that $n = 3$, and your company only has the genres 1 (*action*), 2 (*comedy*) and 3 (*horror*), and the following four movies:
- `movie1`, *Steve of the Dead*, is a horror and a comedy,
- `movie2`, *World War Y*, is a horror and action movie,
- `movie3`, *Lethal Weapon 20*, is a comedy and action movie, and
- `movie4`, *Dumb and Dumbest*, is a comedy movie.

Write the genre vectors $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$ for these movies.

(c) With the same set up of movies as in part (b), suppose a particular user has rated `movie1` a 5, has rated `movie2` a 3, and has rated *movie3* a 4. What rating would this user give to `movie4`?

4. *Orthogonal projections.* Suppose $\mathbf{v} \in \mathbb{R}^n$ is a fixed non-zero vector. If $\mathbf{x} \in \mathbb{R}^n$, the *orthogonal projection of* $\mathbf{x}$ *onto* $\mathbf{v}$ is defined to be the vector

$$\mathbf{Proj_v}(\mathbf{x}) := \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}.$$

(note that since $\mathbf{v} \neq \mathbf{0}$, the denominator $\mathbf{v}\cdot\mathbf{v} = ||\mathbf{v}||^2 \neq 0$, so the above definition makes sense). By properties of the dot product (see also Levandosky pp. 92-93), two useful facts about projection are:
- If θ is the angle between $\mathbf{x}$ and $\mathbf{v}$, then $||\mathbf{Proj_v}(\mathbf{x})|| = ||\mathbf{x}|||\cos\theta|$; and
- The difference vector $\mathbf{x} - \mathbf{Proj_v}(\mathbf{x})$ is orthogonal to $\mathbf{v}$.

(a) Use the formula above to compute $\mathbf{Proj_v}(\mathbf{x})$ and $\mathbf{x} - \mathbf{Proj_v}(\mathbf{x})$ when $\mathbf{x}$ is orthogonal to $\mathbf{v}$.

(b) Use the formula above to compute $\mathbf{Proj_v}(\mathbf{x})$ and $\mathbf{x} - \mathbf{Proj_v}(\mathbf{x})$ when $\mathbf{x} = c\mathbf{v}$ for some scalar c.

(c) For this problem and parts (d)-(e), let $n = 2$, $\mathbf{v} = \left[\begin{smallmatrix}6\\2\end{smallmatrix}\right]$ and $\mathbf{x} = \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right]$. First, compute the vectors $\mathbf{Proj_v}(\mathbf{x})$ and $\mathbf{x} - \mathbf{Proj_v}(\mathbf{x})$.

(d) Letting $\mathbf{v}$ and $\mathbf{x}$ be in part (c), on a diagram in the xy-place,
- draw and label all three vectors $\mathbf{x}$, $\mathbf{v}$, and $\mathbf{Proj_v}(\mathbf{x})$ in standard position, and
- draw and label the vector $\mathbf{x} - \mathbf{Proj_v}(\mathbf{x})$ as the third side of a triangle involving some of the other vectors you have drawn.

Explain why the terms *orthogonal* and *projection* each make sense for $\mathbf{Proj_v}(\mathbf{x})$ here.

(e) What is the shortest distance from the point $(2, 4)$ to the line $y = \frac{1}{3}x$? (*Hint*: think about your diagram in part (d), and recall a fact from geometry about shortest distances.