Linear Algebra (MATH 3333 — 04) Spring 2011

Bonus Project—Ranking Sports Teams

Due: Fri. May. 13, 5pm

Instructions: Here is a project for which you almost certainly *must* use a computer or calculator. Pick your favorite sport and conference. You may consult me or other classmates for help. If you choose, you may work with up to one other classmate and submit as a joint project (provide both students actually work together), or simply do it individually.

Value: Up to 3 final percentage points per individual.

1. Gather the win-loss records for each game between two teams in the conference (different sports have different ways of splitting teams into divisions, but pick something where you are comparing at the minimum 5 or 6 teams). (If they play a lot of games per season in your sport you can restrict yourself to the first 10–20 games for each team. If possible, try and choose this so each team has played each other team about the same number of times. What is more important however is the following: you must use enough games so that there are no two groups of teams in the conference such that no team in one group has played no team in the other group.)

2. Form a win-loss point differential matrix as follows: Let w_{ij} be the sum of all the score differentials (number of points won by) for each game team i beat team j. (If team i never beat team j, $w_{ij} = 0$.) Let W be the $n \times n$ matrix $[w_{ij}]$ where n is the number of teams in your conference.

For example suppose there are only three teams: Team 1, Team 2 and Team 3. Each week two teams play, and the results are as follows:

Week 1: Team 1 beats Team 2 by 5
Week 2: Team 2 beats Team 3 by 3
Week 3: Team 1 beats Team 3 by 2
Weak 4: Team 1 beats Team 2 by 2
Week 5: Team 3 beats Team 2 by 4
Week 6: Team 3 beats Team 1 by 1

Then $w_{12} = 7$ since Team 1 beat Team 2 by 5 in Week 1 and 2 in Week 4. Also $w_{13} = 2$ from the Week 3 result. In summary we will get,

$$W = \begin{pmatrix} 0 & 7 & 2 \\ 0 & 0 & 3 \\ 1 & 4 & 0 \end{pmatrix}.$$

3. Form the normalized matrix A by scaling each column of W so that the sum of the entries of that column is 1. For our example:

$$A = \begin{pmatrix} 0 & \frac{7}{11} & \frac{2}{5} \\ 0 & 0 & \frac{3}{5} \\ 1 & \frac{4}{11} & 0 \end{pmatrix}.$$

4. Rank the teams using the same method as the basic Google Pagerank algorithm (see Section 8.3)—i.e., find a dominant eigenvector (an eigenvector with maximum possible eigenvalue) and use the entries to rank each team. Three remarks: First, for techical reasons, we don't want any row of all 0's in A (these will be for teams who have not won a single game). If there is such a team, delete the corresponding row (and column) from the matrix and put that team in last place. Second, it is possible that the dominant eigenspace has dimension > 1. This will be very rare, but if you notice it happens for your matrix, try slightly changing it. (In the actual Pagerank algorithm, the matrix is tweaked so that neither of these two technicalities, nor the issue in the last sentence of 1, become a problem.)

Second, there are two ways to compute the dominant eigenvector. One is to find the dominant eigenvalue, by determining all eigenvalues, and solve for an eigenvector. The second is to compute the limit of $A^t v_0$ for some random v_0 as $t \to \infty$ (cf. discussion after Theorem 8.5 on p. 504). You may do whichever is easier for you. You are encouraged to use a computer package such as Mathematica (which is free through the IT Store) for these computations.

5. Compare these rankings with the official regular season rankings as well as post/end-season results (if applicable—for simplicity you're only ranking one conference here, which doesn't predict any playoff results for games with teams from a different conference).

6. Write up your results/findings with an explanation of how you obtained them and printouts of any computer calculations. You may submit this for bonus credit to me either at the final exam or by leaving it under my office door (PHSC 924) by 5pm Friday May. 13th.