Sports Ranking Using a Minimum Violations Method

By: Hannah Olejniczak, Bryan Jacobs, Katie Kujawa

Objective

What is Minimum Violations Ranking?

A ranking method that uses pairwise comparisons of n items, that aims to reduce the number of inconsistencies when ranking those items.

Application

The data we chose to analyze is the 2023 SUNYAC Men's Soccer teams. We will apply this to find the final rankings at the end of regular season play.

Point Differential Matrix

We will use a point differential matrix to produce a ranking of our teams.

However we need to rearrange the rows and columns of the matrix in order to apply our MVR.

$$
\begin{array}{c}
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
9 & 0 & 4 & 0 & 2 \\
5 & 0 & 0 & 0 & 0 \\
15 & 3 & 8 & 0 & 5 \\
6 & 0 & 3 & 0 & 0
\end{pmatrix}
\end{array}
$$

Hillside Matrix

A matrix is considered to be in hillside form if it has both of the following:

1. Ascending order across rows
2. Descending order down columns

$$\begin{array}{cc} & \begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array} \\ \begin{array}{c}1\\2\\3\\4\\5\end{array} & \begin{pmatrix} 0 & 3 & 5 & 8 & 15 \\ 0 & 0 & 2 & 4 & 9 \\ 0 & 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\end{array}$$

Violations

$$\begin{array}{cc} & \begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array} \\ \begin{array}{c}1\\2\\3\\4\\5\end{array} & \begin{pmatrix} 0 & 3 & 5 & 8 & 15 \\ 0 & 0 & 2 & 4 & 9 \\ 0 & 0 & 0 & 3 & 6 \\ 0 & 3 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\end{array}$$

Upsets

$$\begin{array}{cc} & \begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array} \\ \begin{array}{c}1\\2\\3\\4\\5\end{array} & \begin{pmatrix} 0 & 3 & 5 & 2 & 15 \\ 0 & 0 & 2 & 4 & 9 \\ 0 & 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\end{array}$$

Weak Wins

City Plot

A common way to visualize entries in a matrix

- Allows for easier recognition of trends
- Bar height represents number of goals team *x* beat team *y* by

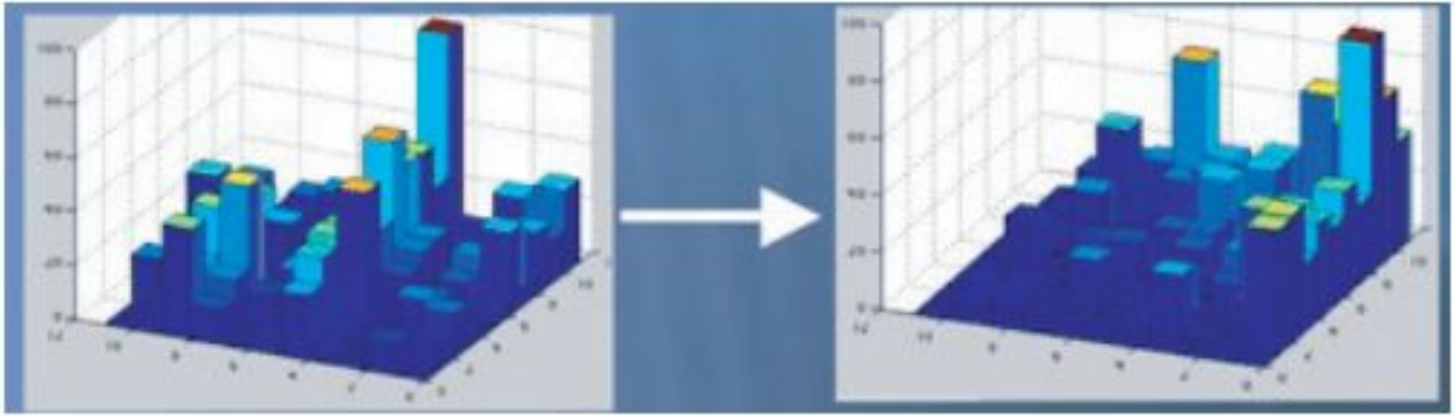


Fig. 1 Cityplot of 11 × 11 data matrix with original ordering and MVR reordering

Why this MVR method is better than others

This MVR method allows for both input and output ties while very few methods allow for either. The new features this method contributes is

(1) Produces an optimal ranking that minimizes upsets and weak win violations

(2) Could include ties

(3) Identifies alternative optimal rankings if they exist

(4) Compares each team's performance against every other team in the league

Formulating our problem as an optimization

The formula behind the reordering of the D matrix to hillside form. The Q matrix is swapping the rows and columns to put them in the correct order.

The Q matrices rows are made up variations of the identity (all 0s with one 1) and the position of the one decides where the rows and columns of D will switch.

$$\min_{\mathbf{Q}} \quad \#\text{ hillside violations of } \mathbf{Q}^T\mathbf{D}\mathbf{Q}$$
$$s.t. \quad \mathbf{Q}^T\mathbf{e} = \mathbf{e}$$
$$\mathbf{e}^T\mathbf{Q} = \mathbf{e}^T$$
$$q_{ij} \in \{0, 1\}$$

The optimization problem above has linear constraints, binary variables and a quadratic objective function which make it difficult to solve as an optimization problem.

Reformulating the problem as a Binary Integer Linear Programming problem allows for an easier solve.

Binary Integer Linear Programming

Determines instances for which one team outperforms another. These instances are compiled to form the actual ranking.

Components needed:

D Matrix: Point differential matrix

C Matrix: - The matrix of constants that is used to compute the number of violations to hillside form
- calculated from the D
- matrix using the algorithm on next slide

X Matrix: the matrix of decision variables that determine if team *i* should be ranked above team *j*

D matrix

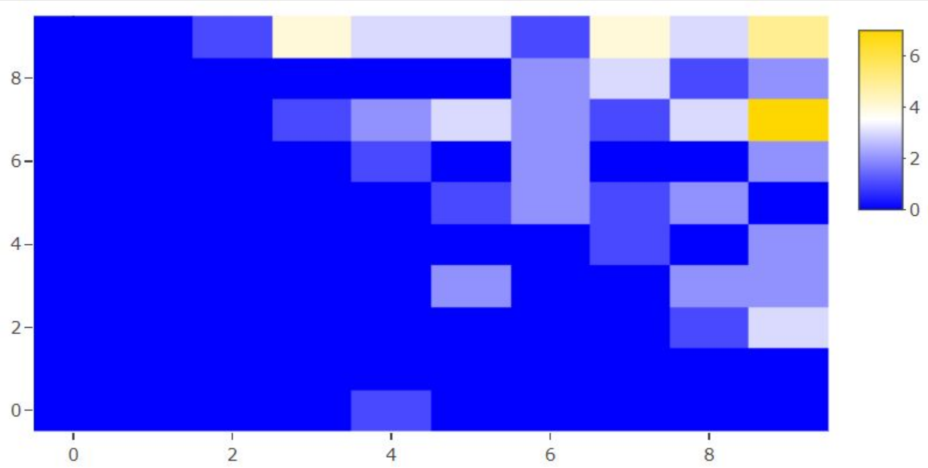
	Geneseo	Oneonta	Cortland	Oswego	Plattsburgh	Brockport	Potsdam	Buff State	Fredonia	New Paltz
Geneseo	-	0	1	0	1	0	3	0	2	0
Oneonta	1	-	0	3	0	0	4	0	1	1
Cortland	0	3	-	2	2	0	5	2	7	2
Oswego	0	0	0	-	0	0	1	0	0	0
Plattsburgh	0	0	0	0	-	0	4	0	1	0
Brockport	2	1	0	1	0	-	3	2	3	0
Potsdam	0	0	0	0	0	0	-	0	0	0
Buff State	2	0	0	2	2	0	1	-	2	0
Fredonia	0	0	0	0	0	0	1	0	-	0
New Paltz	1	0	0	0	0	0	3	2	3	-

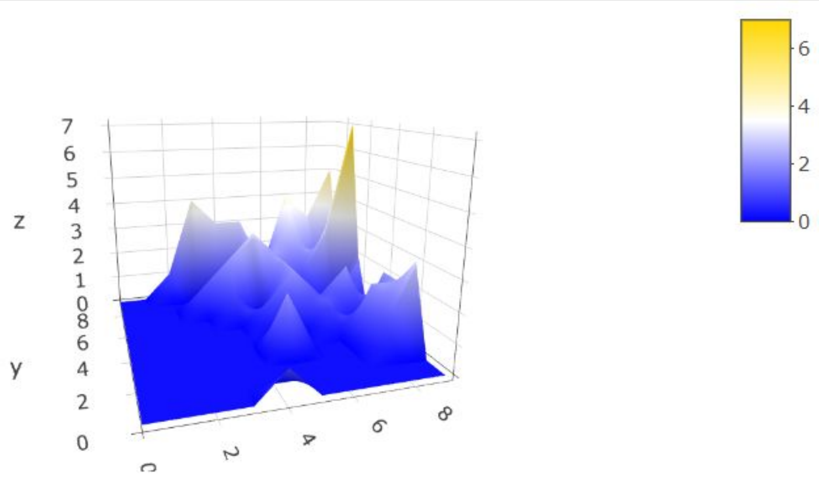
Hillside Form of D

	Cortland	Brockport	Oneonta	Buff State	New Paltz	Geneseo	Plattsburgh	Fredonia	Oswego	Potsdam
Cortland	-	0	3	2	2	0	2	7	2	5
Brockport	0	-	1	2	0	2	0	3	1	3
Oneonta	0	0	-	0	1	1	0	1	3	4
Buff State	0	0	0	-	0	2	2	2	2	1
New Paltz	0	0	0	2	-	1	0	3	0	3
Geneseo	1	0	0	0	0	-	1	2	0	3
Plattsburgh	0	0	0	0	0	0	-	1	0	4
Fredonia	0	0	0	0	0	0	0	-	0	1
Oswego	0	0	0	0	0	0	0	0	-	1
Potsdam	0	0	0	0	0	0	0	0	0	-

Visualizations

Matrix Heat Map





3D Surface Plot

Definition of C matrix

Definition 2 of C matrix: Let $\mathbf{C} = [c_{ij}] \ \forall\, i, j = 1, 2, \ldots, n$ be defined as

$$c_{ij} := \#\{\, k \mid d_{ik} < d_{jk} \,\} + \#\{\, k \mid d_{ki} > d_{kj} \,\},$$

$$\#\{\, k \mid d_{ik} < d_{jk} \,\}$$

Number of teams receiving a lower point differential against team *i* than team *j*

$$\#\{\, k \mid d_{ki} > d_{kj} \,\}$$

Number of team receiving a greater point differential against team *i* than team *j*.

C matrix

	Geneseo	Oneonta	Cortland	Oswego	.Plattsburgh	Brockport	Potsdam	Buff State	Fredonia	New Paltz
Geneseo	0	6	11	2	4	7	1	3	0	4
Oneonta	4	0	6	1	2	4	0	4	0	4
Cortland	0	2	0	1	0	2	0	2	0	2
Oswego	6	5	9	0	4	8	1	4	1	7
Plattsburgh	3	5	7	1	0	8	1	4	0	5
Brockport	2	3	6	0	1	0	1	2	0	0
Potsdam	10	12	14	7	8	13	0	12	4	11
Buff State	4	5	3	2	2	5	0	0	0	3
Fredonia	7	10	12	5	6	11	2	9	0	8
New Paltz	2	2	6	0	2	5	0	4	0	0

Purpose of the X Matrix

The X-matrix determines if team *i* should be ranked above team *j*,

$$x_{ij} = \begin{cases} 1, \text{ if item } i \text{ is ranked above item } j \\ 0, \text{ otherwise.} \end{cases}$$

Once the X-matrix is solved the sum of each row will determine the rankings. To make sure violations remain at a minimum constraints are added to X

$$\begin{aligned} x_{ij} + x_{ji} &= 1 \quad \text{for all distinct pairs } (i, j) \\ x_{ij} + x_{jk} + x_{ki} &\leq 2 \quad \text{for all distinct triples } (i, j, k) \end{aligned}$$

X matrix

	Geneseo	.Oneonta	.Cortland	Oswego	Plattsburgh	Brockport	Potsdam	Buff State	Fredonia	New Paltz
Geneseo	0	0	0	1	0	0	1	1	1	0
Oneonta	1	0	0	1	1	0	1	1	1	0
.Cortland	1	1	0	1	1	1	1	1	1	1
Oswego	0	0	0	0	1	0	1	0	1	0
Plattsburgh	1	0	0	0	0	0	1	0	1	0
Brockport	1	1	0	1	1	0	1	1	1	1
Potsdam	0	0	0	0	0	0	0	0	0	0
Buff State	0	0	0	1	1	0	1	0	1	1
Fredonia	0	0	0	0	0	0	1	0	0	0
New Paltz	1	1	0	1	1	0	1	0	1	0

Final Rankings

Our Ranking:

1. Cortland (9)
2. Brockport (8)
3. Oneonta (6)
4. New Paltz (6)
5. Buffalo State (5)
6. Geneseo (4)
7. Plattsburgh (3)
8. Oswego (3)
9. Fredonia (1)
10. Potsdam (0)

SUNYAC Ranking:

	PTS	CONF	OVERALL
Cortland - 1	22	7-1-1	16-3-3
Brockport - 2	21	6-0-3	9-3-5
Oneonta - 3	17	5-2-2	15-4-3
Buffalo State - 4	16	5-3-1	11-5-3
New Paltz - 5	15	4-2-3	9-3-4
Geneseo - 6	13	4-4-1	9-7-3
Plattsburgh - E	10	2-3-4	7-3-6
Oswego - E	7	1-4-4	2-6-7
Fredonia - E	4	1-7-1	5-8-4
Potsdam - E	0	0-9	4-11-1

Multiple Optimal Solutions

When adding up across the rows of the X matrix, if multiple teams have the same sum, this indicates that there are multiple optimal solutions to the problem.

To verify that a solution still satisfies the objective function, you can swap neighboring rows and columns and see what happens to the hillside form of matrix X.

For example looking at our ranking in X,
Cortland and brockport:
$C(1,2) \neq C(2,1)$

Oneonta and New Paltz:
$C(3,4)=C(4,3)$

If we had more time...

- Solve a larger scale, more significant problem
- Carry out MVR method using coding software
  - Create D, C, and X matrices
- Explore constraint relaxation and sensitivity analysis further
- Compare MVR rankings with official rankings, and perhaps other ranking methods

Thank you.

Hannah Olejniczak, Bryan Jacobs, Katie Kujawa