Bayesian preference learning with the Mallows ranking model

Marta Crispino
Mistis team, Inria Grenoble
marta.crispino@inria.fr

Bayesian Statistics in the Big Data Era
CIRM, November, 26th, 2018
Joint work with

Øystein Sørensen  
University of Oslo

Valeria Vitelli  
University of Oslo

Elja Arjas  
University of Helsinki

Arnoldo Frigessi  
University of Oslo
Preference data is everywhere

- customers express preferences about products and services;
- users select movies on an internet platform (e.g., Netflix);
- genes are ordered based on their expression levels under various experimental conditions.
Preference data is everywhere

- customers express preferences about products and services;
- users select movies on an internet platform (e.g., Netflix);
- genes are ordered based on their expression levels under various experimental conditions.

A ranking represents a statement about the relative quality or relevance of the items being ranked: taste, adherence to a specific user profile, relevance to the biological process under investigation.
Motivation

Preference data is everywhere

- customers express preferences about products and services;
- users select movies on an internet platform (e.g., Netflix);
- genes are ordered based on their expression levels under various experimental conditions.

A ranking represents a statement about the relative quality or relevance of the items being ranked: taste, adherence to a specific user profile, relevance to the biological process under investigation.

Assessors rank items: as panels, users, patients.
Ingredients for Ranking data

A set of **items**, to be evaluated...
Ingredients for Ranking data

A set of **items**, to be evaluated...

...and a pool of **assessors** to evaluate them
Ingredients for Ranking data

A set of **items**, to be evaluated... ...and a pool of **assessors** to evaluate them

A **ranking** is simply a linear ordering of the items
Types of ranking data

**FULL RANKINGS**

<table>
<thead>
<tr>
<th>User</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A10</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>User 2</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>User 3</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>User 4</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>User 5</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>10</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>User 6</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>User 7</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>User 8</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>User 9</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>User 10</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>User 11</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>User 12</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>User 13</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>User 14</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>User 15</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

**PARTIAL RANKINGS**

(top-k)

<table>
<thead>
<tr>
<th>User</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A10</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>User 2</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>User 3</td>
<td>3</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>User 4</td>
<td>4</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>User 5</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>User 6</td>
<td>3</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>User 7</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>User 8</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>User 9</td>
<td>2</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>User 10</td>
<td>NA</td>
<td>3</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>User 11</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>User 12</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>User 13</td>
<td>3</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>4</td>
</tr>
<tr>
<td>User 14</td>
<td>3</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>User 15</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

**PAIRWISE COMPARISONS**

<table>
<thead>
<tr>
<th>User 1</th>
<th>{A3 &lt; A5, A7 &lt; A5}</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 2</td>
<td>{A2 &lt; A9, A6 &lt; A5, A6 &lt; A10, A8 &lt; A1, A8 &lt; A7}</td>
</tr>
<tr>
<td>User 3</td>
<td>{A1 &lt; A9, A4 &lt; A5, A4 &lt; A10, A8 &lt; A7, A9 &lt; A2}</td>
</tr>
<tr>
<td>User 4</td>
<td>{A1 &lt; A4, A2 &lt; A9, A3 &lt; A4, A7 &lt; A4, A9 &lt; A1}</td>
</tr>
<tr>
<td>User 5</td>
<td>{A4 &lt; A3, A4 &lt; A7, A7 &lt; A3, A7 &lt; A10}</td>
</tr>
<tr>
<td>User 6</td>
<td>{A2 &lt; A8, A1 &lt; A2, A8 &lt; A1}</td>
</tr>
<tr>
<td>User 7</td>
<td>{A4 &lt; A1, A9 &lt; A3, A10 &lt; A5}</td>
</tr>
<tr>
<td>User 8</td>
<td>{A2 &lt; A4, A8 &lt; A4, A9 &lt; A5}</td>
</tr>
<tr>
<td>User 9</td>
<td>{A1 &lt; A7, A5 &lt; A9, A10 &lt; A8, A10 &lt; A9}</td>
</tr>
<tr>
<td>User 10</td>
<td>{A1 &lt; A10, A2 &lt; A4, A3 &lt; A4, A3 &lt; A5}</td>
</tr>
<tr>
<td>User 11</td>
<td>{A1 &lt; A8, A9 &lt; A6}</td>
</tr>
<tr>
<td>User 12</td>
<td>{A1 &lt; A5, A7 &lt; A5, A8 &lt; A7, A9 &lt; A7, A10 &lt; A3}</td>
</tr>
<tr>
<td>User 13</td>
<td>{A2 &lt; A10, A4 &lt; A7, A4 &lt; A9, A6 &lt; A3, A6 &lt; A5}</td>
</tr>
<tr>
<td>User 14</td>
<td>{A1 &lt; A4, A1 &lt; A9}</td>
</tr>
<tr>
<td>User 15</td>
<td>{A2 &lt; A8, A3 &lt; A10, A5 &lt; A6, A7 &lt; A8, A9 &lt; A1}</td>
</tr>
</tbody>
</table>
Typical statistical problems

FULL RANKINGS

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A10</th>
</tr>
</thead>
<tbody>
<tr>
<td>user 1</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>user 2</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>user 3</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>user 4</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>user 5</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>10</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>user 6</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>user 7</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>user 8</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>user 9</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>user 10</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>user 11</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>user 12</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>user 13</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>user 14</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>user 15</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

PARTIAL RANKINGS (top-k)

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A10</th>
</tr>
</thead>
<tbody>
<tr>
<td>user 1</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>user 2</td>
<td>NA</td>
<td>3</td>
<td>4</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>user 3</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>user 4</td>
<td>4</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>user 5</td>
<td>3</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>user 6</td>
<td>3</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>user 7</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>user 8</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
</tr>
<tr>
<td>user 9</td>
<td>2</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
</tr>
<tr>
<td>user 10</td>
<td>NA</td>
<td>3</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>4</td>
<td>2</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>user 11</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>user 12</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>user 13</td>
<td>3</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>user 14</td>
<td>3</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>user 15</td>
<td>NA</td>
<td>4</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
</tbody>
</table>

PAIRWISE COMPARISONS

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>A7</th>
<th>A8</th>
<th>A9</th>
<th>A10</th>
</tr>
</thead>
<tbody>
<tr>
<td>user 1</td>
<td>(A3 &lt; A5), (A7 &lt; A5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 2</td>
<td>(A2 &lt; A9), (A6 &lt; A5), (A6 &lt; A10), (A8 &lt; A1), (A8 &lt; A7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 3</td>
<td>(A1 &lt; A9), (A4 &lt; A5), (A4 &lt; A10), (A8 &lt; A7), (A9 &lt; A2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 4</td>
<td>(A1 &lt; A4), (A2 &lt; A9), (A3 &lt; A4), (A7 &lt; A4), (A9 &lt; A1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 5</td>
<td>(A4 &lt; A3), (A4 &lt; A7), (A7 &lt; A3), (A7 &lt; A10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 6</td>
<td>(A2 &lt; A8), (A1 &lt; A2), (A8 &lt; A1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 7</td>
<td>(A4 &lt; A1), (A9 &lt; A3), (A10 &lt; A5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 8</td>
<td>(A2 &lt; A4), (A8 &lt; A4), (A9 &lt; A5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 9</td>
<td>(A1 &lt; A7), (A5 &lt; A9), (A10 &lt; A4), (A10 &lt; A8), (A10 &lt; A9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 10</td>
<td>(A1 &lt; A10), (A2 &lt; A4), (A3 &lt; A4), (A3 &lt; A5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 11</td>
<td>(A1 &lt; A8), (A9 &lt; A6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 12</td>
<td>(A1 &lt; A5), (A7 &lt; A5), (A8 &lt; A7), (A9 &lt; A7), (A10 &lt; A3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 13</td>
<td>(A2 &lt; A10), (A4 &lt; A7), (A4 &lt; A9), (A6 &lt; A3), (A6 &lt; A5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 14</td>
<td>(A1 &lt; A4), (A1 &lt; A9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user 15</td>
<td>(A2 &lt; A8), (A3 &lt; A10), (A5 &lt; A6), (A7 &lt; A8), (A9 &lt; A1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
General setting of the Mallows model

- Let \( \mathcal{P}_n \), be the space of \( n \)-dim permutations
- A ranking, \( R = (R_1, ..., R_n) \), of \( n \) labelled items \( A = \{A_1, ..., A_n\} \) is an element of \( \mathcal{P}_n \), where, for all \( i \), \( R_i \) is the rank assigned to item \( A_i \).

\[
\begin{array}{cccccccccc}
A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} \\
e.g. & 1 & 7 & 8 & 2 & 10 & 4 & 6 & 9 & 3 & 5
\end{array}
\]
General setting of the Mallows model

- Let $\mathcal{P}_n$, be the space of $n$-dim permutations
- A ranking, $R = (R_1, ..., R_n)$, of $n$ labelled items $A = \{A_1, ..., A_n\}$ is an element of $\mathcal{P}_n$, where, for all $i$, $R_i$ is the rank assigned to item $A_i$.

$$
\begin{array}{cccccccccc}
A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} \\
e.g. & 1 & 7 & 8 & 2 & 10 & 4 & 6 & 9 & 3 \end{array}
$$

- The Mallows model (Mallows, 1957) gives the probability density for $R \in \mathcal{P}_n$,

$$
P(R \mid \alpha, \rho) := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right]
$$

- $\rho \in \mathcal{P}_n$: location parameter, shared **consensus ranking**
- $d(\cdot, \cdot)$: right-invariant (Diaconis, 1988) distance between permutations (example)
- $\alpha \geq 0$: scale parameter
- $Z_n(\alpha)$: partition function
General setting of the Mallows model

- Let $\mathcal{P}_n$, be the space of $n$-dim permutations.
- A ranking, $R = (R_1, ..., R_n)$, of $n$ labelled items $A = \{A_1, ..., A_n\}$ is an element of $\mathcal{P}_n$, where, for all $i$, $R_i$ is the rank assigned to item $A_i$.

\[ R = (1, 7, 8, 2, 10, 4, 6, 9, 3, 5) \]

- The Mallows model (Mallows, 1957) gives the probability density for $R \in \mathcal{P}_n$,

\[
P(R | \alpha, \rho) := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right]
\]

- $\rho \in \mathcal{P}_n$: location parameter, shared consensus ranking
- $d(\cdot, \cdot)$: right-invariant (Diaconis, 1988) distance between permutations (example)
- $\alpha \geq 0$: scale parameter
- $Z_n(\alpha)$: partition function

- **Flexibility in the choice of the distance** (driven by the application), (example)
  - Cayley, Hamming, Ulam: measures of disorder $\rightarrow$ genomics, cryptography
  - Footrule ($l_1$), Spearman ($l_2$), Kendall: domain of preferences $\rightarrow$ elections, movies
The Mallows model: challenge

\[ P(R \mid \alpha, \rho) := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right] \]

Challenge for inference: computation of the partition function

\[ Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} \exp \left[ -\frac{\alpha}{n} d(r, 1_n) \right] \]
The Mallows model: challenge

\[ P(R | \alpha, \rho) := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right] \]

Challenge for inference: computation of the partition function

\[ Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} \exp \left[ -\frac{\alpha}{n} d(r, 1_n) \right] \]

With Kendall, Cayley and Hamming distances → \( Z_n(\alpha) \) has closed form
The Mallows model: challenge

\[ \begin{align*}
P(R | \alpha, \rho) & := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right] \\
\end{align*} \]

Challenge for inference: computation of the partition function

\[ Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} \exp \left[ -\frac{\alpha}{n} d(r, 1_n) \right] \]

With Kendall, Cayley and Hamming distances \( \rightarrow \) \( Z_n(\alpha) \) has closed form

With footrule and Spearman (i.e. \( l_1 \) and \( l_2 \)) distances \( \rightarrow \) no results.
So far solved numerically for very small values of \( n \), as infeasible for larger \( n \).
The Mallows model: challenge

\[ P(R \mid \alpha, \rho) := \frac{1}{Z_n(\alpha)} \exp \left[ -\frac{\alpha}{n} d(R, \rho) \right] \]

Challenge for inference: computation of the partition function

\[ Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} \exp \left[ -\frac{\alpha}{n} d(r, 1_n) \right] \]

With Kendall, Cayley and Hamming distances → \( Z_n(\alpha) \) has closed form

With \textbf{footrule and Spearman} (i.e. \( l_1 \) and \( l_2 \)) distances → no results.
So far solved numerically for very small values of \( n \), as infeasible for larger \( n \).

Our approach:

1. Strategy to compute \( Z_n(\alpha) \) \underline{exactly} for moderately large values of \( n \).
2. When needed for larger \( n \), Importance Sampling (IS) scheme.
Bayesian inference: full rankings

- $N$ users rank $n$ items $A = \{A_1, \ldots, A_n\}$
- Data $R = \{R_j\}_{j=1}^N \rightarrow \text{full rankings}$
- $R_j = (R_{j1}, \ldots, R_{jn}) \in \mathcal{P}_n$: ranking given by user $j$ to the full set of items
- $R_{ji}$: rank given to item $A_i$ by user $j$. 
Bayesian inference: full rankings

- \( N \) users rank \( n \) items \( \mathcal{A} = \{A_1, \ldots, A_n\} \)
- Data \( \mathbf{R} = \{\mathbf{R}_j\}_{j=1}^N \rightarrow \) full rankings
- \( \mathbf{R}_j = (R_{j1}, ..., R_{jn}) \in \mathcal{P}_n \): ranking given by user \( j \) to the full set of items
- \( R_{ji} \): rank given to item \( A_i \) by user \( j \).
- Statistical model: \( \mathbf{R}_1, \ldots, \mathbf{R}_N | \alpha, \rho \overset{i.i.d.}{\sim} \text{Mallows}(\alpha, \rho) \)

\[
P(\mathbf{R}_1, \ldots, \mathbf{R}_N; \alpha, \rho) = \frac{1}{Z_n(\alpha)^N} \exp \left\{ -\frac{\alpha}{n} \sum_{j=1}^N d(\mathbf{R}_j, \rho) \right\}
\]

- Prior: assume independence between \( \rho \) and \( \alpha \) and no prior information
  - \( \rho \): uniform over \( \mathcal{P}_n \rightarrow \pi(\rho) = \frac{1}{n!} \mathcal{P}_n(\rho) \)
  - \( \alpha \): (truncated) exponential prior
- Posterior density

\[
\pi(\rho, \alpha | \mathbf{R}_1, \ldots, \mathbf{R}_N) \propto \frac{1}{Z_n(\alpha)^N} \exp \left\{ -\alpha \left[ n^{-1} \sum_{j=1}^N d(\mathbf{R}_j, \rho) + \lambda \right] \right\}
\]
Bayesian inference: top$-k$ rankings

- $N$ users rank a - **possibly different** - subset of items $A_j \subseteq \{A_1, A_2, \ldots, A_n\}$
- Typical situation: Each user only assesses her **top$-k_j$ preferred items**
- Data $R = \{R_j\}_{j=1}^{N} \rightarrow$ partial rankings
Bayesian inference: top–$k$ rankings

- $N$ users rank a possibly different subset of items $A_j \subseteq \{A_1, A_2, \ldots, A_n\}$
- Typical situation: Each user only assesses her top–$k_j$ preferred items
- Data $R = \{R_j\}_{j=1}^N \rightarrow$ partial rankings

Apply **data augmentation** techniques: estimating the lacking ranks consistently with the partial observations.

- Define augmented full rankings $\tilde{R}_1, \ldots, \tilde{R}_N$, where each $\tilde{R}_j$ is compatible with the partial informations in $R_j$

*Example*:

<table>
<thead>
<tr>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
<th>$R_4$</th>
<th>$R_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>NA</td>
</tr>
</tbody>
</table>

- Posterior density

$$
\pi(\alpha, \rho|R_1, \ldots, R_N) = \sum_{\tilde{R}_1 \in S_1} \cdots \sum_{\tilde{R}_N \in S_N} P(\alpha, \rho, \tilde{R}_1, \ldots, \tilde{R}_N|R_1, \ldots, R_N).
$$

where $S_j$, set of rankings compatible with $R_j$, $j = 1, \ldots, N$. 

Marta Crispino
Bayesian inference: transitive pair comparisons

- $N$ users do not see all the possible items, but only express binary preferences between pairs of them.
- Data $\{B_j\}_{j=1}^N$ are sets of pair preferences, of the form $(A_{m_1} \prec A_{m_2})$ if $A_{m_1}$ preferred to $A_{m_2}$.
Bayesian inference: transitive pair comparisons

- $N$ users do not see all the possible items, but only express binary preferences between pairs of them.
- Data $\{B_j\}_{j=1}^N$ are sets of pair preferences, of the form $(A_{m_1} < A_{m_2})$ if $A_{m_1}$ preferred to $A_{m_2}$.
- Define augmented full rankings $\tilde{R}_1, \ldots, \tilde{R}_N$, where each $\tilde{R}_j$ is compatible with the partial informations in (the transitive closure of) $B_j$.

\begin{align*}
\tilde{R}_1 &= (5, 4, 2, 1, 3) \\
\tilde{R}_1 &= (2, 5, 3, 1, 4) \\
\tilde{R}_1 &= (3, 5, 2, 1, 4) \\
\tilde{R}_1 &= (4, 5, 2, 1, 3)
\end{align*}

- Posterior density

$$
\pi(\alpha, \rho|B_1, \ldots, B_N) = \sum_{\tilde{R}_1 \in \text{tc}(B_1)} \cdots \sum_{\tilde{R}_N \in \text{tc}(B_N)} P(\alpha, \rho|\tilde{R}_1, \ldots, \tilde{R}_N).
$$
Bayesian inference: non-transitive pair comparisons

- Same setting as before BUT users allowed to be inconsistent in their choices
- E.g. It may occur a non-transitive pattern in the data

\[ B_j = \{ A_5 < A_2, A_2 < A_3, A_3 < A_5, \ldots \} \]

- Ideally we want to "coherentize" the preferences, and estimate the latent truth.
- Idea: assume non-transitive patterns arise because of mistakes made by the users
- Identification/correction of mistakes: borrowing strength
Bayesian inference: non-transitive pair comparisons

- **Posterior density**

\[
\pi(\alpha, \rho|\mathcal{B}_1, \ldots, \mathcal{B}_N) = \sum_{\tilde{R}_1 \in \mathcal{P}_n} \ldots \sum_{\tilde{R}_N \in \mathcal{P}_n} P(\alpha, \rho|\tilde{R}_1, \ldots, \tilde{R}_N) P(\tilde{R}_1, \ldots, \tilde{R}_N|\mathcal{B}_1, \ldots, \mathcal{B}_N)
\]

- **Assumption:**

\[
P(\tilde{R}_1, \ldots, \tilde{R}_N|\mathcal{B}_1, \ldots, \mathcal{B}_N) = \prod_{j=1}^{N} P(\tilde{R}_j|\mathcal{B}_j)
\]

- **P(\tilde{R}_j|\mathcal{B}_j):** Weight of each full rank in the sum

- **Interpretation:** probability of ordering the pairs as in \(\mathcal{B}_j\) when the latent ranking for user \(j\) is \(\tilde{R}_j\) → probability of making mistakes in the binary choices
  - **Random mistake:** independent of the pair of items
  - **Logistic model:** the likelihood of a mistake increases if the items are perceived as similar by the user  
    (details)
Implementation and examples

**Implementation**: Metropolis within Gibbs MCMC, with data augmentation

Many **applications** (require mixture extension):

- **Sushi** benchmark data: full rankings, heterogeneity (*
- **Meta analysis of gene expression** data: partial rankings (*
- Preference among **beach pictures**: pairwise comparisons (*
- **Sound Data**: pairwise comparisons with many non-transitive patterns, due to difficult perception, heterogeneity (*
- **Movie preferences**: very sparse pairwise comparison data, comparison with Collaborative Filtering (*
Conclusions

- Ongoing work

  - **R package** BayesMallows, available on CRAN
  - **Conjugate prior** for $\rho$ (joint work with I. Antoniano-Villalobos) (idea)
  - **Genomics** application: Mixture of Mallows for detection of differential gene expression (joint work with V. Djordjilovic)

Thanks for your attention!
Conclusions

- Ongoing work
  - R package BayesMallows, available on CRAN
  - Conjugate prior for $\rho$ (joint work with I. Antoniano-Villalobos) (idea)
  - Genomics application: Mixture of Mallows for detection of differential gene expression (joint work with V. Djordjilovic)

- Future
  - Extension to rankings with ties (to model indifference in the preference)
  - Integration of covariates (of items and/or of users)
  - Variable selection: rank only the items which are worth being ranked
  - Un-equal quality of assessors
Conclusions

● Ongoing work

  o R package BayesMallows, available on CRAN

  o Conjugate prior for $\rho$ (joint work with I. Antoniano-Villalobos) (idea)

  o Genomics application: Mixture of Mallows for detection of differential gene expression (joint work with V. Djordjilovic)

● Future

  o Extension to rankings with ties (to model indifference in the preference)

  o Integration of covariates (of items and/or of users)

  o Variable selection: rank only the items which are worth being ranked

  o Un-equal quality of assessors

Thanks for your attention!


Right-invariance

Definition: Right-invariant distance

A distance function is right-invariant, if \( d(\rho_1, \rho_2) = d(\rho_1 \eta, \rho_2 \eta) \) for all \( \eta, \rho_1, \rho_2 \in \mathcal{P}_n \), where \( \rho \eta = \rho \circ \eta = \rho \eta = (\rho_{\eta_1}, \ldots, \rho_{\eta_n}) \).

Example

- 4 students, \((A_1, A_2, A_3, A_4)\), admitted in a PhD program
- initial ranking \( \rho_1 = (1, 3, 4, 2) \) (admission)
- final ranking \( \rho_2 = (3, 4, 1, 2) \) (general exam)
- \( d(\rho_1, \rho_2) \) can be thought of as a measure of the goodness of judgement of the PhD admission board.
- If the students are relabelled in a different ordering, for example \((A_4, A_2, A_1, A_3)\), then \( \rho_1 \eta = (2, 3, 1, 4) \) and \( \rho_2 \eta = (2, 4, 3, 1) \), where \( \eta = (4, 2, 1, 3) \) determines the relabelling of the students.
- Natural to assume \( d(\rho_1, \rho_2) = d(\rho_1 \eta, \rho_2 \eta) \), because the situation depicted is the same.

\[
\begin{array}{cccc}
A_1 & A_2 & A_3 & A_4 \\
\hline
\rho_1 & 1 & 3 & 4 & 2 \\
\rho_2 & 3 & 4 & 1 & 2 \\
\end{array}
\quad \rightarrow \quad
\begin{array}{cccc}
A_4 & A_2 & A_1 & A_3 \\
\hline
\rho_1 \eta & 2 & 3 & 1 & 4 \\
\rho_2 \eta & 2 & 4 & 3 & 1 \\
\end{array}
\]
Right-invariance

Consequence of right-invariance
For any $\rho_1, \rho_2 \in \mathcal{P}_n$, it holds $d(\rho_1, \rho_2) = d(\rho_1 \rho_2^{-1}, 1_n)$, where $1_n = (1, 2, \ldots, n)$. Then $Z_n(\alpha, \rho)$ is free of $\rho$, as

$$Z_n(\alpha, \rho) = \sum_{r \in \mathcal{P}_n} e^{-\alpha_n d(r, \rho)} = \sum_{r \in \mathcal{P}_n} e^{-\alpha_n d(r \rho^{-1}, 1_n)} = \sum_{r' \in \mathcal{P}_n} e^{-\alpha_n d(r', 1_n)} = Z_n(\alpha)$$

Common right-invariant distances between permutations $\rho_1, \rho_2 \in \mathcal{P}_n$
- Footrule ($l_1$): $d_F(\rho_1, \rho_2) = \sum_{i=1}^n |\rho_{1i} - \rho_{2i}|$
- Spearman ($l_2$): $d_S(\rho_1, \rho_2) = \sum_{i=1}^n (\rho_{1i} - \rho_{2i})^2$
- Kendall: minimum number of adjacent transpositions which convert $\rho_1$ into $\rho_2$
- Cayley: minimum number of transpositions which convert $\rho_1$ into $\rho_2$
- Ulam: minimum number of deletion-insertion operations to convert $\rho_1$ into $\rho_2$.
- Hamming: minimum number of substitutions required to convert $\rho_1$ into $\rho_2$.

Go back
Distances: why important

Consider the following two permutations:

\[ \sigma = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \]
\[ \tau = (9, 10, 3, 4, 5, 6, 7, 8, 1, 2) \]

First and second elements of \( \sigma \), are at the bottom of \( \tau \).
Distances: why important

Consider the following two permutations:

\[ \sigma = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \]
\[ \tau = (9, 10, 3, 4, 5, 6, 7, 8, 1, 2) \]

First and second elements of \( \sigma \), are at the bottom of \( \tau \).

If \( \sigma \) and \( \tau \) represent preferences about movies \( \rightarrow \) very different profiles.
Distances: why important

Consider the following two permutations:

\[ \sigma = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \]
\[ \tau = (9, 10, 3, 4, 5, 6, 7, 8, 1, 2) \]

First and second elements of \( \sigma \), are at the bottom of \( \tau \).

If \( \sigma \) and \( \tau \) represent preferences about movies \( \rightarrow \) very different profiles.

If \( \sigma \) and \( \tau \) represent genomes \( \rightarrow \) just one translocation in the genome
Distances: why important

Consider the following two permutations:

\[ \sigma = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \]
\[ \tau = (9, 10, 3, 4, 5, 6, 7, 8, 1, 2) \]

First and second elements of \( \sigma \), are at the bottom of \( \tau \).

If \( \sigma \) and \( \tau \) represent preferences about movies \( \rightarrow \) very different profiles.

If \( \sigma \) and \( \tau \) represent genomes \( \rightarrow \) just one translocation in the genome.

Normalized Spearman \((l_2)\): \( d_S(\sigma, \tau) \approx 0.5 \)
Normalized Cayley: \( d_C(\sigma, \tau) \approx 0.28 \)
The Mallows density

\[ \rho = 1_n = (1, 2, \ldots, n) \]

\[ \frac{1}{n!} = \frac{1}{120} \]

- Go back
Exact computation of $Z_n(\alpha)$

$Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} e^{-\frac{\alpha}{n} d(r, 1_n)} = \sum_{d_i \in \mathcal{D}} |L_i| e^{-\frac{\alpha}{n} d_i}$

where

- $d(r, 1_n) \in \mathcal{D} = \{d_1, ..., d_h\}$, $h$ depends on $n$ and $d(\cdot, \cdot)$
- $L_i = \{r \in \mathcal{P}_n : d(r, 1_n) = d_i\} \subset \mathcal{P}_n$, $i = 1, ..., h$. 
Exact computation of $Z_n(\alpha)$

$$Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} e^{-\frac{\alpha}{n} d(r, 1_n)} =$$

$$= \sum_{d_i \in \mathcal{D}} |L_i| e^{-\frac{\alpha}{n} d_i}$$

where

- $d(r, 1_n) \in \mathcal{D} = \{d_1, \ldots, d_h\}$, $h$ depends on $n$ and $d(\cdot, \cdot)$
- $L_i = \{r \in \mathcal{P}_n : d(r, 1_n) = d_i\} \subset \mathcal{P}_n$, $i = 1, \ldots, h$.

Sufficient to know $|L_i|$, for all values $d_i \in \mathcal{D} \rightarrow$ Easier, but still unfeasible for large $n$
Exact computation of $Z_n(\alpha)$

\[
Z_n(\alpha) = \sum_{r \in \mathcal{P}_n} e^{-\frac{\alpha}{n} d(r, 1_n)} = \\
\sum_{d_i \in \mathcal{D}} \left| L_i \right| e^{-\frac{\alpha}{n} d_i}
\]

where

- $d(r, 1_n) \in \mathcal{D} = \{d_1, \ldots, d_h\}$, $h$ depends on $n$ and $d(\cdot, \cdot)$
- $L_i = \{r \in \mathcal{P}_n : d(r, 1_n) = d_i\} \subset \mathcal{P}_n$, $i = 1, \ldots, h$.

Sufficient to know $|L_i|$, for all values $d_i \in \mathcal{D} \rightarrow$ Easier, but still unfeasible for large $n$

**Special cases solution** (from the computer programming field)

- **Footrule distance**: $\mathcal{D} = \{0, 2, 4, \ldots, \lfloor n^2/2 \rfloor\}$, $|L_i|$ is the sequence A062869 tabulated for $n \leq 50$ in the On-Line Encyclopedia of Integer Sequences (OEIS)
- **Spearman’s distance**: $\mathcal{D}\{0, 2, 4, \ldots, 2\left(\frac{n+1}{3}\right)\}$, $|L_i|$ is the sequence A175929 tabulated only until $n \leq 14$ in the OEIS
Importance Sampling approximation of $Z_n(\alpha)$

Let $R^1, \ldots, R^K$ sampled from auxiliary distribution $q(R)$, then

$$\hat{Z}_n(\alpha) = K^{-1} \sum_{k=1}^{K} \exp \left[-(\alpha/n)d(R^k, 1_n)\right] q(R^k)^{-1}.$$  

**Pseudo-likelihood approach:** Let $\{i_1, \ldots, i_n\}$ be a uniform sample from $\mathcal{P}_n$, giving the order of the pseudo-likelihood factorization. Then

$$P(R_{in} | 1_n) = \frac{\exp \left[-(\alpha/n)d(R_{in}, i_n)\right] \cdot 1_{[1, \ldots, n]}(R_{in})}{\sum_{r_n \in \{1, \ldots, n\}} \exp \left[-(\alpha/n)d(r_n, i_n)\right]} ,$$

$$P \left( R_{i_{n-1}} | R_{in}, 1_n \right) = \frac{\exp \left[-(\alpha/n)d(R_{i_{n-1}}, i_{n-1})\right] \cdot 1_{\{1, \ldots, n\}\setminus\{R_{in}\}}(R_{i_{n-1}})}{\sum_{r_{n-1} \in \{1, \ldots, n\}\setminus\{R_{in}\}} \exp \left[-(\alpha/n)d(r_{n-1}, i_{n-1})\right]} ,$$

$$\vdots$$

$$P \left( R_{i_2} | R_{i_3}, \ldots, R_{in}, 1_n \right) = \frac{\exp \left[-(\alpha/n)d(R_{i_2}, i_2)\right] \cdot 1_{\{1, \ldots, n\}\setminus\{R_{i_3}, \ldots, R_{in}\}}(R_{i_2})}{\sum_{r_2 \in \{1, \ldots, n\}\setminus\{R_{i_3}, \ldots, R_{in}\}} \exp \left[-(\alpha/n)d(r_2, i_2)\right]} ,$$

$$P \left( R_{i_1} | R_{i_2}, \ldots, R_{in}, 1_n \right) = 1_{\{1, \ldots, n\}\setminus\{R_{i_2}, \ldots, R_{in}\}}(R_{i_1}).$$
IS approximation of $Z_n(\alpha)$

Example: $n = 50$, footrule distance

Mukherjee (2016) limit: asymptotic approximation of $Z_n(\alpha)$
Effect of the approximation of $Z_n(\alpha)$ on inference
Consider a sample of rankings \( R_1, \ldots, R_N \mid \rho, \theta \sim \mathcal{M}_S(\theta, \rho) \), where \( \mathcal{M}_S(\cdot, \cdot) \) is the Mallows density with \( \theta = \alpha / n \), and Spearman \((l_2)\) distance,

\[
d(\rho, \sigma) = \sum_{i=1}^{n} (\rho_i - \sigma_i)^2
\]

Assume \( \theta \) known, then

\[
P(R_1, \ldots, R_N; \theta, \rho) = \prod_{j=1}^{N} \frac{1}{Z(\theta)} \exp \left\{ -\theta \sum_{i=1}^{n} (R_i - \rho_i)^2 \right\} \propto \exp \left\{ 2\theta N \sum_{i=1}^{n} \rho_i \bar{R}_i \right\},
\]

where \( \bar{R}_i = \frac{1}{N} \sum_{j=1}^{N} R_{ji}, \ i = 1, \ldots, n, \) is the sample average of the \( i \)-th rank.

**Proposition**

Let \( \mathbb{PP}_n \) be the \( n \)-dim permutation polytope, that is, the convex hull of the elements of \( \mathcal{P}_n \). Then \( \bar{R} = (\bar{R}_1, \ldots, \bar{R}_n) \in \mathbb{PP}_n \).
Conjugate prior for $\rho$ (joint work with I. Antoniano-Villalobos)

Keeping $\theta$ fixed, the conjugate prior for $\rho \in \mathcal{P}_n$ is

$$
\pi(\rho|\rho_0, \theta_0) = \frac{1}{Z^*(\theta_0, \rho_0)} \exp \left[ -\theta_0 \sum_{i=1}^{n} (\rho_{0i} - \rho_i)^2 \right] \mathbb{1}(\rho_0 \in \mathcal{P}_n) \mathbb{1}(\theta_0 \in \mathbb{R}^+) \\
\propto \exp \left[ 2\theta_0 \sum_{i=1}^{n} \rho_i \rho_{0i} \right]
$$

The posterior density for $\rho$ is

$$
\pi(\rho|R_1, \ldots, R_N) \propto \exp \left\{ 2(\theta_0 + \theta N) \sum_{i=1}^{n} \rho_i \left[ \frac{\theta N}{\theta_0 + \theta N} \bar{R}_i + \frac{\theta_0}{\theta_0 + \theta N} \rho_{0i} \right] \right\}
$$

i.e. $\pi(\rho|R_1, \ldots, R_N)$ same parametric density of the prior, with updated parameters

$$
\rho_N = \frac{\theta N}{\theta_0 + \theta N} \bar{R} + \frac{\theta_0}{\theta_0 + \theta N} \rho_0 \\
\theta_N = \theta_0 + \theta N
$$

The result reminds Diaconis and Ylvisaker (1979)
Conjugate prior for $\rho$ (joint work with I. Antoniano-Villalobos)

Example: $n = 3$, $N = 40$, $\theta = 0.5$, $\rho = (3, 2, 1)$.
Sample and obtain $\bar{R} = (2.25, 2.125, 1.625)$.

$\rho_0 = (1, 2, 3)$, varying $\theta_0 = 0, 10, 20, 30$.  

$\rho_0 = (1, 2.5, 2.5)$, varying $\theta_0 = 0, 10, 20, 30$.  

\[ \rho \]
\[ \rho_0 \]

Go back
Non-transitive pairwise preferences

- **Mouse click mistake:**

  \[ P(\text{mistake} \mid \theta, R_j) = \theta, \quad \theta \in [0, 0.5) \]

- **Logistic model**

  \[
  \text{logit} \ P(\text{mistake} \mid R_j, \beta_0, \beta_1) = -\beta_0 - \beta_1 \frac{d_{R_j,m}}{n-1}
  \]

  where \( d_{R_j,m} = |R_{j1} - R_{j2}| \) if \( B_{j,m} = (O_1 < O_2) \).
How important is 3-D spatial motion to our understanding of human agency?

- $n = 12$ abstract sounds, made from the action of a cellist while playing, each obtained starting at the best representation of the original gesture, and then reducing or removing some aspects of the sound

  SOUND1
  Full sonification, the best one can make to capture motion - based on what we know about our perception and hearing

  SOUND7
  Like the previous one, with pitch modulation removed

  SOUND10
  The ‘worst’ sonification, spatial variation is flattened, both pitch and volume variations removed.
A group of $N = 46$ listeners repeatedly presented with pairs of sounds and asked to choose the one that most evokes the sense of human causation (or physicality)

To what extent listeners report non-transitive sets of preferences?
A group of $N = 46$ listeners repeatedly presented with pairs of sounds and asked to choose the one that most evokes the sense of human causation (or physicality).

To what extent listeners report non-transitive sets of preferences? The percentage of listeners who report at least one non-transitivity is 80%.
A group of $N = 46$ listeners repeatedly presented with pairs of sounds and asked to choose the one that most evokes the sense of human causation (or physicality).

To what extent listeners report non-transitive sets of preferences? The percentage of listeners who report at least one non-transitivity is 80%.

We expect the listeners to be clustered: differences in the interpretation of the test and in how people listen to sounds $\rightarrow$ Mixture model generalization of the main model.
Sound data: non-transitive pair comparisons (with N. Barrett)

Posterior consensus ranking $\rho$ of the 3 clusters

Expert explanation of the clusters:
- Cluster 1: listeners who like slower spatial variation
- Cluster 2: listeners who are listening spatially
- Cluster 3: negative preference for spatial motion
Posterior probabilities for all the sonifications of being ranked among the top-4 for the 3 clusters
Sound data: non-transitive pair comparisons (with N. Barrett)

Probability that the best sonified sounds are amongst the top-4 ranked sounds (obtained thanks to the estimated individual rankings).
SAA: index measuring listeners’ awareness of spatial audio (3 is highly aware)

Spatial listening is a skill that is enhanced through training
Sushi data: full rankings

\[ N = 5000 \] Japanese people interviewed: each gives his/her complete ranking of \( n = 10 \) sushi variants (items)
$N = 5000$ Japanese people interviewed: each gives his/her complete ranking of $n = 10$ sushi variants (items)
Sushi data: full rankings

MAP estimate

<table>
<thead>
<tr>
<th></th>
<th>$c = 1$</th>
<th>$c = 2$</th>
<th>$c = 3$</th>
<th>$c = 4$</th>
<th>$c = 5$</th>
<th>$c = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_c$</td>
<td>0.243 (0.23,0.26)</td>
<td>0.131 (0.12,0.14)</td>
<td>0.107 (0.11,0.12)</td>
<td>0.117 (0.11,0.12)</td>
<td>0.121 (0.11,0.13)</td>
<td>0.278 (0.27,0.29)</td>
</tr>
<tr>
<td>$\alpha_c$</td>
<td>3.62 (3.52,3.75)</td>
<td>2.55 (2.35,2.71)</td>
<td>3.8 (3.42,4.06)</td>
<td>4.02 (3.78,4.26)</td>
<td>4.46 (4.25,4.68)</td>
<td>1.86 (1.77,1.94)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>Item</th>
<th>Item</th>
<th>Item</th>
<th>Item</th>
<th>Item</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fatty tuna</td>
<td>shrimp</td>
<td>sea urchin</td>
<td>fatty tuna</td>
<td>fatty tuna</td>
<td>fatty tuna</td>
</tr>
<tr>
<td>2</td>
<td>sea urchin</td>
<td>sea eel</td>
<td>egg</td>
<td>shrimp</td>
<td>salmon roe</td>
<td>tuna</td>
</tr>
<tr>
<td>3</td>
<td>salmon roe</td>
<td>squid</td>
<td>squid</td>
<td>squid</td>
<td>squid</td>
<td>salmon roe</td>
</tr>
<tr>
<td>4</td>
<td>sea eel</td>
<td>tuna</td>
<td>tuna roll</td>
<td>tuna roll</td>
<td>salmon roe</td>
<td>tuna</td>
</tr>
<tr>
<td>5</td>
<td>tuna</td>
<td>shrimp</td>
<td>squid</td>
<td>squid</td>
<td>sea eel</td>
<td>salmon roe</td>
</tr>
<tr>
<td>6</td>
<td>shrimp</td>
<td>squid</td>
<td>tuna roll</td>
<td>tuna roll</td>
<td>egg</td>
<td>squids</td>
</tr>
<tr>
<td>7</td>
<td>squid</td>
<td>tuna roll</td>
<td>tuna roll</td>
<td>salmon roe</td>
<td>cucumber roll</td>
<td>squid</td>
</tr>
<tr>
<td>8</td>
<td>tuna roll</td>
<td>fatty tuna</td>
<td>cucumber roll</td>
<td>egg</td>
<td>sea eel</td>
<td>salmon roe</td>
</tr>
<tr>
<td>9</td>
<td>egg</td>
<td>salmon roe</td>
<td>cucumber roll</td>
<td>egg</td>
<td>sea eel</td>
<td>sea urchin</td>
</tr>
<tr>
<td>10</td>
<td>cucumber roll</td>
<td>sea urchin</td>
<td>sea urchin</td>
<td>sea urchin</td>
<td>sea urchin</td>
<td>sea urchin</td>
</tr>
</tbody>
</table>


» Go back
Beaches data: pairwise comparisons

- $n = 15$ images of tropical beaches shown in pairs to $N = 60$ users (25 random pairs each)
- Question: “Which of the two beaches would you prefer to go to in your next vacation?”
Beaches data: pairwise comparisons

- $n = 15$ images of tropical beaches shown in pairs to $N = 60$ users (25 random pairs each)
- Question: “Which of the two beaches would you prefer to go to in your next vacation?”
Beaches data: pairwise comparisons

- We can also estimate the individual rankings

![Heatmap of pairwise comparisons between beaches and users]

Go back
Meta-analysis in Genomics: top–k rankings

Context:
- Studies of differential gene expression between two conditions produce a list of genes, ranked according to their level of differential expression as measured by some test statistics.
Context:

- Studies of differential gene expression between two conditions produce a list of genes, ranked according to their level of differential expression as measured by some test statistics.
- Little agreement among gene lists found by independent studies comparing the same conditions leads to difficulties in finding a consensus list over all available studies. This situation raises the question of whether a consensus top list over all available studies can be found.
Meta-analysis in Genomics: top–$k$ rankings

Context:

- Studies of differential gene expression between two conditions produce a list of genes, ranked according to their level of differential expression as measured by some test statistics.
- Little agreement among gene lists found by independent studies comparing the same conditions leads to difficulties in finding a consensus list over all available studies. This situation raises the question of whether a consensus top list over all available studies can be found.
- Biologists are often concerned with the few most relevant genes in the specific context of the pathology, to set in place further more detailed lab experiments.
- $N = 5$ studies comparing prostate cancer patients with healthy controls, based on differential gene expression
- Each study produces top–25 (i.e. $k = 25$) list of genes (unique genes $n = 89$)
The fact that \( n \gg N \), and having partial data, both contribute to keeping precision small.

However, the posterior probability for each gene to be among the top-10 or top-25 is not so low, thus demonstrating that our approach can provide a valid criterion for consensus (with uncertainty quantification).
Bayesian Mallows model VS Collaborative Filtering (with Q. Liu)

BMM

CF

number of correct predictions for the next 2 items

\( \lambda \)

Go back