Can Rawls’ ideas of fairness be embodied in $k$-means clustering?

It seems so!

"... they [social and economic inequalities] are to be to the greatest benefit of the least-advantaged members of society."

– Rawls (Justice as Fairness)

John Rawls (1921–2002) was an influential 20th-century moral and political philosopher in liberal tradition. He is frequently cited in courts of law and by politicians in the US and UK. His ideas of fairness are regarded as time-tested and a good mix of pragmatism and principledness.

1. $k$-means Clustering

Unlabelled Data

Labelled Clusters

image taken from shorturl.at/osyEW

2. Fairness Notion

Do centroids represent groups ‘fairly’?

<table>
<thead>
<tr>
<th>(Popular notion) Demographic Parity</th>
<th>Rawlsian Fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum inequality</td>
<td>allow inequality, but represent least-advantaged group better</td>
</tr>
</tbody>
</table>

3. Research Questions

Is there a cluster assignment satisfying Rawlsian Fairness?

Yes, approximately

indicated by ● in the plot below

Can $k$-means clusters be perturbed to embody Rawlsian Fairness?

Yes, but can be improved

trajectories R1 and R2 moving from ● towards ● in the plot below