# Estimating Bilateral Migration

Guy J. Abel

# IPFP

- A common problem with bilateral migration data is that it is unavailable or outdated.
  - Often collected in censuses
- In some cases there are other data sources available that provide information on the in- and out-migration totals
  - Population registers

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | 100 | 30 | 70 | 200 |
| B | 50 | | 45 | 5 | 100 |
| C | 60 | 35 | | 40 | 135 |
| D | 20 | 25 | 20 | | 65 |
| Sum | 130 | 160 | 95 | 115 | 500 |

*B. Primary Data: Marginal Totals*

| *Origin* | *Destination* | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | | | | 250 |
| B | | | | | 75 |
| C | | | | | 125 |
| D | | | | | 150 |
| Sum | 150 | 200 | 50 | 200 | 600 |

# IPFP

- Data in the above situation, where the marginal tables totals are known but the cell values are unknown, can be estimated using a range of methods
- Similar data estimation challenges exist for more detailed migration flow tables, for example:
  - In- and out-migration totals by age in each region are known, but the origin-destination migration flow table for each age group is missing.
  - Required by multi-regional cohort-component models
  - Estimating international migration flows from stocks (see for example Abel (2013) )

# IPFP

- A popular approach to estimate values in a contingency table based on known marginal tables and an initial contingency table is the
- First described by Deming and Stephan (1940), the IPFP has since been widely studied in a number of different disciplines and under a number of synonyms such as raking, matrix scaling or the RAS algorithm
  - Lovelace et al. (2015) gives a good overview of the application of IPFP in social sciences.
  - Lomax and Norman (2016) for another overview more specific to demography.
- Mathematical approach to iteratively adjust a *seed* contingency table $\mu_{ij}^{(0)} = m_{ij}$ to known row and column totals ($n_{i+}$ and $n_{+j}$)

$$\mu_{ij}^{(t+1)} = \frac{\mu_{ij}^{(t)}}{\mu_{i+}^{(t)}} n_{i+} \qquad \mu_{ij}^{(t+2)} = \frac{\mu_{ij}^{(t+1)}}{\mu_{+j}^{(t+1)}} n_{+j}$$

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | *Totals* |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | 100 | 30 | 70 | 200 |
| B | 50 | | 45 | 5 | 100 |
| C | 60 | 35 | | 40 | 135 |
| D | 20 | 25 | 20 | | 65 |
| Sum | 130 | 160 | 95 | 115 | 500 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | *Totals* |
|---|---|---|---|---|---|
| | A | B | C | D | Sum | In |
| A | | 100 | 30 | 70 | 200 | 130 |
| B | 50 | | 45 | 5 | 100 | 160 |
| C | 60 | 35 | | 40 | 135 | 95 |
| D | 20 | 25 | 20 | | 65 | 115 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | *Totals* | |
|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out |
| A | | 100 | 30 | 70 | 200 | 130 | 200 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | | | | | | | | | |
| B | | | | | | | | | | |
| C | | | | | | | | | | |
| D | | | | | | | | | | |
| Sum | | | | | | | | | | |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | | | | | 150 | 250 | | | |
| B | | | | | | 200 | 75 | | | |
| C | | | | | | 50 | 125 | | | |
| D | | | | | | 200 | 150 | | | |
| Sum | | | | | | 600 | 600 | | | |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 150 | 250 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 200 | 75 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 50 | 125 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 200 | 150 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 600 | 600 | 500 | 500 | 0 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 125.00 | 37.50 | 87.50 | 250.00 | 150 | 250 | 139.21 | 250.00 | -110.79 |
| B | 37.50 | | 33.75 | 3.75 | 75.00 | 200 | 75 | 215.10 | 75.00 | 140.10 |
| C | 55.56 | 32.41 | | 37.04 | 125.00 | 50 | 125 | 117.40 | 125.00 | -7.60 |
| D | 46.15 | 57.69 | 46.15 | | 150.00 | 200 | 150 | 128.29 | 150.00 | -21.71 |
| Sum | 139.21 | 215.10 | 117.40 | 128.29 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 116.23 | 15.97 | 136.41 | 268.61 | 150 | 250 | 150.00 | 268.61 | -118.61 |
| B | 40.41 | | 14.37 | 5.85 | 60.63 | 200 | 75 | 200.00 | 60.63 | 139.37 |
| C | 59.86 | 30.13 | | 57.74 | 147.74 | 50 | 125 | 50.00 | 147.74 | -97.74 |
| D | 49.73 | 53.64 | 19.66 | | 123.03 | 200 | 150 | 200.00 | 123.03 | 76.97 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 108.17 | 14.86 | 126.96 | 250.00 | 150 | 250 | 161.27 | 250.00 | -88.73 |
| B | 49.99 | | 17.78 | 7.23 | 75.00 | 200 | 75 | 199.07 | 75.00 | 124.07 |
| C | 50.65 | 25.50 | | 48.86 | 125.00 | 50 | 125 | 56.61 | 125.00 | -68.39 |
| D | 60.63 | 65.40 | 23.96 | | 150.00 | 200 | 150 | 183.05 | 150.00 | 33.05 |
| Sum | 161.27 | 199.07 | 56.61 | 183.05 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 108.68 | 13.13 | 138.72 | 260.53 | 150 | 250 | 150.00 | 260.53 | -110.53 |
| B | 46.49 | | 15.70 | 7.90 | 70.10 | 200 | 75 | 200.00 | 70.10 | 129.90 |
| C | 47.11 | 25.61 | | 53.38 | 126.10 | 50 | 125 | 50.00 | 126.10 | -76.10 |
| D | 56.40 | 65.71 | 21.17 | | 143.27 | 200 | 150 | 200.00 | 143.27 | 56.73 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 104.29 | 12.60 | 133.11 | 250.00 | 150 | 250 | 155.49 | 250.00 | -94.51 |
| B | 49.74 | | 16.80 | 8.45 | 75.00 | 200 | 75 | 198.47 | 75.00 | 123.47 |
| C | 46.70 | 25.39 | | 52.91 | 125.00 | 50 | 125 | 51.56 | 125.00 | -73.44 |
| D | 59.05 | 68.79 | 22.16 | | 150.00 | 200 | 150 | 194.48 | 150.00 | 44.48 |
| Sum | 155.49 | 198.47 | 51.56 | 194.48 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 105.09 | 12.22 | 136.89 | 254.20 | 150 | 250 | 150.00 | 254.20 | -104.20 |
| B | 47.99 | | 16.29 | 8.69 | 72.98 | 200 | 75 | 200.00 | 72.98 | 127.02 |
| C | 45.05 | 25.59 | | 54.41 | 125.05 | 50 | 125 | 50.00 | 125.05 | -75.05 |
| D | 56.96 | 69.32 | 21.49 | | 147.78 | 200 | 150 | 200.00 | 147.78 | 52.22 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 103.35 | 12.01 | 134.63 | 250.00 | 150 | 250 | 152.17 | 250.00 | -97.83 |
| B | 49.32 | | 16.75 | 8.94 | 75.00 | 200 | 75 | 199.30 | 75.00 | 124.30 |
| C | 45.03 | 25.58 | | 54.39 | 125.00 | 50 | 125 | 50.57 | 125.00 | -74.43 |
| D | 57.82 | 70.37 | 21.81 | | 150.00 | 200 | 150 | 197.96 | 150.00 | 47.96 |
| Sum | 152.17 | 199.30 | 50.57 | 197.96 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 103.72 | 11.88 | 136.02 | 251.62 | 150 | 250 | 150.00 | 251.62 | -101.62 |
| B | 48.62 | | 16.56 | 9.03 | 74.20 | 200 | 75 | 200.00 | 74.20 | 125.80 |
| C | 44.39 | 25.67 | | 54.95 | 125.01 | 50 | 125 | 50.00 | 125.01 | -75.01 |
| D | 56.99 | 70.62 | 21.57 | | 149.18 | 200 | 150 | 200.00 | 149.18 | 50.82 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 103.05 | 11.80 | 135.15 | 250.00 | 150 | 250 | 150.84 | 250.00 | -99.16 |
| B | 49.14 | | 16.73 | 9.13 | 75.00 | 200 | 75 | 199.72 | 75.00 | 124.72 |
| C | 44.39 | 25.66 | | 54.95 | 125.00 | 50 | 125 | 50.22 | 125.00 | -74.78 |
| D | 57.31 | 71.01 | 21.69 | | 150.00 | 200 | 150 | 199.22 | 150.00 | 49.22 |
| Sum | 150.84 | 199.72 | 50.22 | 199.22 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 103.20 | 11.75 | 135.67 | 250.62 | 150 | 250 | 150.00 | 250.62 | -100.62 |
| B | 48.87 | | 16.66 | 9.16 | 74.69 | 200 | 75 | 200.00 | 74.69 | 125.31 |
| C | 44.14 | 25.70 | | 55.16 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.10 | 21.59 | | 149.69 | 200 | 150 | 200.00 | 149.69 | 50.31 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.94 | 11.72 | 135.34 | 250.00 | 150 | 250 | 150.32 | 250.00 | -99.68 |
| B | 49.07 | | 16.73 | 9.20 | 75.00 | 200 | 75 | 199.89 | 75.00 | 124.89 |
| C | 44.14 | 25.70 | | 55.16 | 125.00 | 50 | 125 | 50.09 | 125.00 | -74.91 |
| D | 57.11 | 71.25 | 21.63 | | 150.00 | 200 | 150 | 199.70 | 150.00 | 49.70 |
| Sum | 150.32 | 199.89 | 50.09 | 199.70 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 103.00 | 11.70 | 135.54 | 250.24 | 150 | 250 | 150.00 | 250.24 | -100.24 |
| B | 48.97 | | 16.70 | 9.21 | 74.88 | 200 | 75 | 200.00 | 74.88 | 125.12 |
| C | 44.04 | 25.71 | | 55.25 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.29 | 21.60 | | 149.88 | 200 | 150 | 200.00 | 149.88 | 50.12 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.90 | 11.69 | 135.41 | 250.00 | 150 | 250 | 150.12 | 250.00 | -99.88 |
| B | 49.04 | | 16.73 | 9.23 | 75.00 | 200 | 75 | 199.96 | 75.00 | 124.96 |
| C | 44.04 | 25.71 | | 55.24 | 125.00 | 50 | 125 | 50.03 | 125.00 | -74.97 |
| D | 57.04 | 71.35 | 21.62 | | 150.00 | 200 | 150 | 199.88 | 150.00 | 49.88 |
| Sum | 150.12 | 199.96 | 50.03 | 199.88 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.92 | 11.68 | 135.49 | 250.09 | 150 | 250 | 150.00 | 250.09 | -100.09 |
| B | 49.00 | | 16.72 | 9.23 | 74.95 | 200 | 75 | 200.00 | 74.95 | 125.05 |
| C | 44.01 | 25.72 | | 55.28 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.36 | 21.60 | | 149.95 | 200 | 150 | 200.00 | 149.95 | 50.05 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.88 | 11.68 | 135.44 | 250.00 | 150 | 250 | 150.05 | 250.00 | -99.95 |
| B | 49.03 | | 16.73 | 9.24 | 75.00 | 200 | 75 | 199.98 | 75.00 | 124.98 |
| C | 44.01 | 25.72 | | 55.28 | 125.00 | 50 | 125 | 50.01 | 125.00 | -74.99 |
| D | 57.01 | 71.39 | 21.61 | | 150.00 | 200 | 150 | 199.96 | 150.00 | 49.96 |
| Sum | 150.05 | 199.98 | 50.01 | 199.96 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.89 | 11.67 | 135.47 | 250.04 | 150 | 250 | 150.00 | 250.04 | -100.04 |
| B | 49.02 | | 16.72 | 9.24 | 74.98 | 200 | 75 | 200.00 | 74.98 | 125.02 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.39 | 21.60 | | 149.98 | 200 | 150 | 200.00 | 149.98 | 50.02 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.45 | 250.00 | 150 | 250 | 150.02 | 250.00 | -99.98 |
| B | 49.03 | | 16.73 | 9.24 | 75.00 | 200 | 75 | 199.99 | 75.00 | 124.99 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.40 | 21.60 | | 150.00 | 200 | 150 | 199.98 | 150.00 | 49.98 |
| Sum | 150.02 | 199.99 | 50.00 | 199.98 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.88 | 11.67 | 135.46 | 250.01 | 150 | 250 | 150.00 | 250.01 | -100.01 |
| B | 49.02 | | 16.73 | 9.24 | 74.99 | 200 | 75 | 200.00 | 74.99 | 125.01 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.40 | 21.60 | | 149.99 | 200 | 150 | 200.00 | 149.99 | 50.01 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.00 | 150 | 250 | 150.01 | 250.00 | -99.99 |
| B | 49.03 | | 16.73 | 9.24 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 199.99 | 150.00 | 49.99 |
| Sum | 150.01 | 200.00 | 50.00 | 199.99 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.01 | 150 | 250 | 150.00 | 250.01 | -100.01 |
| B | 49.03 | | 16.73 | 9.24 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 200.00 | 150.00 | 50.00 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.00 | 150 | 250 | 150.00 | 250.00 | -100.00 |
| B | 49.03 | | 16.73 | 9.24 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.99 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 200.00 | 150.00 | 50.00 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.00 | 150 | 250 | 150.00 | 250.00 | -100.00 |
| B | 49.03 | | 16.73 | 9.25 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.98 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 200.00 | 150.00 | 50.00 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.00 | 150 | 250 | 150.00 | 250.00 | -100.00 |
| B | 49.03 | | 16.73 | 9.25 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.98 | 25.72 | | 55.29 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 200.00 | 150.00 | 50.00 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

*A. Auxiliary Data*

| *Origin* | *Destination* | | | | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | Net |
| A | | 100 | 30 | 70 | 200 | 130 | 200 | -70 |
| B | 50 | | 45 | 5 | 100 | 160 | 100 | 60 |
| C | 60 | 35 | | 40 | 135 | 95 | 135 | -40 |
| D | 20 | 25 | 20 | | 65 | 115 | 65 | 50 |
| Sum | 130 | 160 | 95 | 115 | 500 | 500 | 500 | 0 |

*B. Constrained Estimates*

| *Origin* | *Destination* | | | | | *Target* | | *Totals* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | Sum | In | Out | In | Out | Net |
| A | | 102.87 | 11.67 | 135.46 | 250.00 | 150 | 250 | 150.00 | 250.00 | -100.00 |
| B | 49.03 | | 16.73 | 9.25 | 75.00 | 200 | 75 | 200.00 | 75.00 | 125.00 |
| C | 43.98 | 25.72 | | 55.30 | 125.00 | 50 | 125 | 50.00 | 125.00 | -75.00 |
| D | 56.99 | 71.41 | 21.60 | | 150.00 | 200 | 150 | 200.00 | 150.00 | 50.00 |
| Sum | 150.00 | 200.00 | 50.00 | 200.00 | 600.00 | 600 | 600 | 600.00 | 600.00 | 0.00 |

# IPFP

- Willekens (1999) calls the seed data an *auxiliary* table and notes that it should be information on a variables related to migration.
  - Typically past migration flow data
  - Distances or travel costs between the origin-destination pairs have been used where no past data exists
  - Limited testing to see which seeds work best for estimating migration

# mipfp

- The *mipfp* package by Barthélemy and Suesse (2018) implements IPFP in R using the `Ipfp()` function
- Can be used for multi-dimensional marginal constraint problems.
- Three inputs
  - `seed` a matrix of auxiliary data to aid estimation
  - `target.list` a list of dimensions that are being targeted (see next point)
  - `target.data` a list of targets related to `target.list`
- R numbers dimension of arrays with
  - 1 row
  - 2 column
  - 3 table
  - ...
- The `target.list` might involve
  - a single target, e.g. column totals `target.list = list(2)`
  - multiple targets, e.g. row and column totals `target.list = list(1, 2)`
  - sums over cells rather than margins of array, e.g. cells summed over tables `target.list = list(c(1, 2))`

# mipfp

```
> r <- LETTERS[1:4]
> m0 <- matrix(data = c(0, 100, 30, 70,
+                       50, 0, 45, 5,
+                       60, 35, 0, 40,
+                       20, 25, 20, 0),
+              nrow = 4, ncol = 4, byrow = TRUE,
+              dimnames = list(orig = r, dest = r))
> addmargins(m0)
      dest
orig    A   B  C   D Sum
  A     0 100 30  70 200
  B    50   0 45   5 100
  C    60  35  0  40 135
  D    20  25 20   0  65
  Sum 130 160 95 115 500
```

# mipfp

```
> orig_tot <- c(250, 75, 125, 150)
> dest_tot <- c(150, 200, 50, 200)
> names(orig_tot ) <- names(dest_tot) <- r
>
> orig_tot
  A   B   C   D
250  75 125 150
> dest_tot
  A   B   C   D
150 200  50 200
>
> # check sums are equal
> sum(orig_tot)
[1] 600
> sum(dest_tot)
[1] 600
```

# mipfp

```
> library(mipfp)
> Ipfp(seed = m0, target.list = list(1, 2),
+      target.data = list(orig_tot, dest_tot))

Call:
Ipfp(seed = m0, target.list = list(1, 2), target.data = list(orig_tot,
    dest_tot))

Method:  ipfp - convergence:  TRUE

Estimates:
    dest
orig         A          B         C           D
   A  0.00000 102.87046 11.67024 135.459297
   B 49.02778   0.00000 16.72686   9.245364
   C 43.98433  25.72033  0.00000  55.295339
   D 56.98789  71.40921 21.60290   0.000000
```

# mipfp

```
> # save the result
> y0 <- Ipfp(seed = m0, target.list = list(1, 2),
+            target.data = list(orig_tot, dest_tot))
>
> # view with totals
> addmargins(y0$x.hat)
      dest
orig           A          B        C           D Sum
  A      0.00000 102.87046 11.67024 135.459297 250
  B     49.02778   0.00000 16.72686   9.245364  75
  C     43.98433  25.72033  0.00000  55.295339 125
  D     56.98789  71.40921 21.60290   0.000000 150
  Sum 150.00000 200.00000 50.00000 200.000000 600
```

# Three dimensions

*Auxiliary Data - Female*

| *Origin* | *Destination* | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | 80 | 10 | 55 | 145 |
| B | 30 | | 20 | 0 | 50 |
| C | 50 | 15 | | 10 | 75 |
| D | 5 | 20 | 10 | | 35 |
| Sum | 85 | 115 | 40 | 65 | 305 |

*Auxiliary Data - Male*

| *Origin* | *Destination* | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | 20 | 20 | 15 | 55 |
| B | 20 | | 25 | 5 | 50 |
| C | 10 | 20 | | 30 | 60 |
| D | 15 | 5 | 10 | | 30 |
| Sum | 45 | 45 | 55 | 50 | 195 |

*Primary Data*

| *Origin* | *Destination* | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Sum |
| A | | | | | 250 |
| B | | | | | 75 |
| C | | | | | 125 |
| D | | | | | 150 |
| Sum | 150 | 200 | 50 | 200 | 600 |

# IPFP More Complicated Data Situations

- The IPFP can be used for more complex data situations with more than two dimensions.
- Key to using the `Ipfp()` function is setting the inputs for `target.data`.

```
> library(tidyverse)
> d <- expand_grid(orig = r, dest = r, sex = c("female", "male")) %>%
+   mutate(flow = c(0, 0, 80, 20, 10, 20, 55, 15, 30, 20, 0, 0, 20, 25, 0, 5, 50, 1
>
> d
# A tibble: 32 x 4
   orig  dest  sex      flow
   <chr> <chr> <chr>   <dbl>
 1 A     A     female      0
 2 A     A     male        0
 3 A     B     female     80
 4 A     B     male       20
 5 A     C     female     10
 6 A     C     male       20
 7 A     D     female     55
 8 A     D     male       15
 9 B     A     female     30
10 B     A     male       20
# ... with 22 more rows
# i Use `print(n = ...)` to see more rows
```

# Estimating Detailed Bilateral Migration

```
> m1 <- xtabs(formula = flow ~ orig + dest + sex, data = d)
> m1
, , sex = female

    dest
orig  A  B  C  D
   A  0 80 10 55
   B 30  0 20  0
   C 50 15  0 10
   D  5 20 10  0

, , sex = male

    dest
orig  A  B  C  D
   A  0 20 20 15
   B 20  0 25  5
   C 10 20  0 30
   D 15  5 10  0
```

# mipfp

```
> addmargins(m1, margin = c(1,2))
, , sex = female

      dest
orig    A   B   C   D Sum
  A     0  80  10  55 145
  B    30   0  20   0  50
  C    50  15   0  10  75
  D     5  20  10   0  35
  Sum  85 115  40  65 305

, , sex = male

      dest
orig    A   B   C   D Sum
  A     0  20  20  15  55
  B    20   0  25   5  50
  C    10  20   0  30  60
  D    15   5  10   0  30
  Sum  45  45  55  50 195
```

# mipfp

```
> addmargins(m1)[,,sex = "Sum"]
      dest
orig    A   B   C   D Sum
  A     0 100  30  70 200
  B    50   0  45   5 100
  C    60  35   0  40 135
  D    20  25  20   0  65
  Sum 130 160  95 115 500
>
> # use overall in- and out-flow targets from the previous example
> orig_tot
  A   B   C   D
250  75 125 150
> dest_tot
  A   B   C   D
150 200  50 200
```

# mipfp

```
> y1 <- Ipfp(seed = m1, target.list = list(1, 2),
+            target.data = list(orig_tot, dest_tot))
> addmargins(y1$x.hat, margin = c(1,2))
, , sex = female

      dest
orig           A           B           C           D         Sum
  A     0.000000  82.296369   3.890080 106.432305 192.618755
  B    29.416668   0.000000   7.434158   0.000000  36.850826
  C    36.653611  11.022997   0.000000  13.823835  61.500444
  D    14.246971  57.127369  10.801451   0.000000  82.175792
  Sum  80.317251 150.446736  22.125690 120.256140 373.145817

, , sex = male

      dest
orig           A           B           C           D         Sum
  A     0.000000  20.574092   7.780161  29.026992  57.381245
  B    19.611112   0.000000   9.292698   9.245364  38.149174
  C     7.330722  14.697330   0.000000  41.471504  63.499556
  D    42.740914  14.281842  10.801451   0.000000  67.824208
  Sum  69.682749  49.553264  27.874310  79.743860 226.854183
```

# mipfp

```
> # estimates summed over sex
> addmargins(y1$x.hat)[,,sex = "Sum"]
     dest
orig           A           B           C           D         Sum
  A     0.000000 102.870462  11.670241 135.459297 250.000000
  B    49.027781   0.000000  16.726856   9.245364  75.000000
  C    43.984334  25.720327   0.000000  55.295339 125.000000
  D    56.987886  71.409211  21.602903   0.000000 150.000000
  Sum 150.000000 200.000000  50.000000 200.000000 600.000000
> # targets used
> orig_tot
  A   B   C   D
250  75 125 150
> dest_tot
  A   B   C   D
150 200  50 200
```

# mipfp

```
> y1$x.hat %>%
+   as.data.frame.table(responseName = "est") %>%
+   as_tibble()
# A tibble: 32 x 4
   orig  dest  sex      est
   <fct> <fct> <fct>  <dbl>
 1 A     A     female  0
 2 B     A     female 29.4
 3 C     A     female 36.7
 4 D     A     female 14.2
 5 A     B     female 82.3
 6 B     B     female  0
 7 C     B     female 11.0
 8 D     B     female 57.1
 9 A     C     female  3.89
10 B     C     female  7.43
# ... with 22 more rows
# i Use `print(n = ...)` to see more rows
```

# Net constrained origin-destination flows

- Plane (1981) developed a proportional adjustment algorithm for estimating bilateral migration flows to match both
  - Constraints on the net migration of each region
  - Total sum of the bilateral migration flows
- Requires data on
  - Past bilateral migration flows (or any kind of seed data)
  - Current (target) total migration flows (over whole system)
  - Current (target) net migration flows
  - Distance matrix to correspond
- No function in R for the method, although in *migest* package the `cm_net_tot()` function provides a similar set of estimates
  - Unable to incorporate distance matrix

# Net constrained origin-destination flows

```
> addmargins(m0)
      dest
orig    A   B  C   D Sum
  A     0 100 30  70 200
  B    50   0 45   5 100
  C    60  35  0  40 135
  D    20  25 20   0  65
  Sum 130 160 95 115 500
>
> # observed net
> library(migest)
> sum_region(m0)
# A tibble: 4 x 5
  region out_mig in_mig  turn   net
  <chr>    <dbl>  <dbl> <dbl> <dbl>
1 A          200    130   330   -70
2 B          100    160   260    60
3 C          135     95   230   -40
4 D           65    115   180    50
```

# Net constrained origin-destination flows

- Estimate migration flows to match new net migration and grand total.

```
> y1 <- cm_net_tot(net_tot = c(-100, 125, -75, 50), tot = 600,
+                  m = m0, verbose = FALSE)
> addmargins(y1$n)
     dest
orig           A          B        C           D       Sum
  A      0.00000 136.22513 32.93756  79.068944 248.23163
  B     49.88761   0.00000 42.28296   4.833488  97.00406
  C     74.27815  50.62851  0.00000  47.977516 172.88418
  D     24.06590  35.15032 22.66377   0.000000  81.87999
  Sum 148.23166 222.00396 97.88429 131.879947 599.99986
> # estimated net, matches target net
> sum_region(y1$n)
# A tibble: 4 x 5
  region out_mig in_mig  turn    net
  <chr>    <dbl>  <dbl> <dbl>  <dbl>
1 A         248.   148.   396. -100.
2 B          97.0  222.   319.  125.
3 C         173.    97.9  271.  -75.0
4 D          81.9  132.   214.   50.0
```

# Net constrained origin-destination flows

- The requirement on the total sum of the bilateral flow for the algorithm is not realistic.
  - Plane (1981) method not widely adopted
  - In many countries the overall number of migrant flows, that is demographically consistent with natural population change, is typically not known.
  - If the overall number of migrant flows is known, it is typically obtained from a comprehensive population register, and thus bilateral migration or total in- and out-migration flows already exist. If it is the later, can use IPFP approaches.
- In recent years I have been working on a method that constrains only to the net migration totals.
  - Unpublished, work in progress, use at own risk
  - Method is available in the `cm_net()` function in the *migest* package
- Potential uses
  - Update bilateral migration flows from surveys or administrative data to match known demographic consistent net migration totals
  - Estimate bilateral migration flows from known net migration totals using non-migration data as a seed (if no migration flow data available)

# Net constrained origin-destination flows

```
> y2 <- cm_net(net_tot = c(-100, 125, -75, 50), m = m0, verbose = FALSE)
> addmargins(y2$n)
     dest
orig           A          B        C           D        Sum
  A      0.00000 124.97056 27.96585  71.121910 224.05832
  B     40.00942   0.00000 33.56693   4.065067  77.64142
  C     64.36422  46.92119  0.00000  43.597199 154.88260
  D     19.68451  30.74980 18.34980   0.000000  68.78412
  Sum 124.05815 202.64155 79.88258 118.784175 525.36645
> # estimated net matches target net
> sum_region(y2$n)
# A tibble: 4 x 5
  region out_mig in_mig  turn    net
  <chr>    <dbl>  <dbl> <dbl>  <dbl>
1 A         224.   124.   348. -100.
2 B          77.6  203.   280.  125.
3 C         155.    79.9  235.  -75.0
4 D          68.8  119.   188.   50.0
```

# Exercise (ex7.R)

```r
# 0.  a) Load the KOSTAT2022.Rproj file.
#     Run the getwd() below. It should print the directory where the
#     KOSTAT2022.Rproj file is located.
getwd()
#     b) Load the packages used in this exercise
library(tidyverse)
library(mipfp)
##
##
##
# 1. Run the code below to read in the bilateral data in uk_census_2011_tidy.csv
#    from the ONS 2011 British Census
cen11 <- read_csv("./data/uk_census_2011_tidy.csv")
cen11
# 2. Run the code below to read in the bilateral data in
#    uk_nhs_hesa_2018.csv from the British administrative data (National
#    Health Service patient records and Higher Education Statistics Authority)
nhs18 <- read_csv("./data/uk_nhs_hesa_2018_tidy.csv")
nhs18
# 3. Run the code below to create data with abbreviated region names - to make
#    it easier to view the matrices for each time period
#    Note: the census data is more detailed (orig - dest - age - sex) than the
#          administrative data (orig - dest)
cen11 <- cen11 %>%
  mutate(orig_full = orig,
```

# References I


Abel, Guy J. 2013. "Estimating global migration flow tables using place of birth data." *Demographic Research* 28 (March): 505–46. https://doi.org/10.4054/DemRes.2013.28.18.

Barthélemy, Johan, and Thomas Suesse. 2018. "mipfp : An R Package for Multidimensional Array Fitting and Simulating Multivariate Bernoulli Distributions." *Journal of Statistical Software* 86 (Code Snippet 2). https://doi.org/10.18637/jss.v086.c02.

Deming, W. Edwards, and Frederick F Stephan. 1940. "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known." *The Annals of Mathematical Statistics* 11 (4): 427–44. https://doi.org/10.1214/aoms/1177731829.

Lomax, Nik, and Paul Norman. 2016. "Estimating population attribute values in a table: 'Get me started in' iterative proportional fitting." *Professional Geographer* 68 (3): 451–61. https://doi.org/10.1080/00330124.2015.1099449.

Lovelace, Robin, Mark Birkin, Dimitris Ballas, and Eveline van Leeuwen. 2015. "Evaluating the Performance of Iterative Proportional Fitting for Spatial Microsimulation: New Tests for an Established Technique." *Journal of Artificial Societies and Social Simulation* 18 (2): 1–15. https://doi.org/10.18564/jasss.2768.

Plane, David A. 1981. "Estimation of Place-to-Place Migration Flows from Net Migration Totals: A Minimum Information Approach." *International Regional Science Review* 6 (1): 33–51. https://doi.org/10.1177/016001768100600103.

Willekens, Frans. 1999. "Modeling approaches to the indirect estimation of migration flows: from entropy to EM." *Mathematical Population Studies* 7 (3): 239–78, 308. https://doi.org/10.1080/08898489909525459.