CIE4801 Transportation and spatial modelling
Trip distribution

Rob van Nes, Transport & Planning
17/4/13
Content

• What’s it about

• Three methods
  • With special attention for the first

• Special issues
1.

Trip distribution: what’s it about
Introduction to trip distribution

- Zonal data
- Transport networks
- Travel resistances
- Trip production / Trip attraction
- Trip distribution
- Modal split
- Period of day
- Assignment
- Trip frequency choice
- Destination choice
- Mode choice
- Time choice
- Route choice
- Travel times network loads etc.
What do we want to know?
Introduction to trip distribution

**Given:** Productions and attractions for each zone (i.e. departures and arrivals)

**Determine:** The number of trips from each zone to all other zones

From: zone 1 ...
zone i ...
...

To: zone 1 ...
zone j ...
...

*total production*

*total attraction*
Introduction to trip distribution
Trip distribution: 3 methods

- Analogy based
- Extrapolation or growth
- Choice modelling
2.1

Method 1: Analogy
The gravity model

\[ G_{ij} = g \cdot m_i \cdot m_j \cdot \frac{1}{d_{ij}^2} \]

- \( G_{ij} \) = gravitational force between \( i \) and \( j \)
- \( g \) = gravitational constant
- \( m_i, m_j \) = mass of planet \( i \) (\( j \) respectively)
- \( d_{ij} \) = distance between \( i \) and \( j \)
The gravity model

Assumptions:

Number of trips between an origin and a destination zone is proportional to:
• a production ability factor for the origin zone
• an attraction ability factor for the destination zone
• a factor depending on the travel costs between the zones

Mathematical formulation:

\[ T_{ij} = \rho Q_i X_j F_{ij} \]

- \( T_{ij} \) = # trips from zone \( i \) to zone \( j \)
- \( \rho \) = measure of average trip intensity
- \( Q_i \) = production potential of zone \( i \)
- \( X_j \) = attraction potential of zone \( j \)
- \( F_{ij} \) = accessibility of \( j \) from \( i \)

Possible interpretations of \( Q_i \) and \( X_j \): populations, production & attraction, ...
Trip distribution using the gravity model

Depending on the amount of information available, different models result

- Production and attraction are both unknown
  - Direct demand model

- Only production or attraction is known
  - Singly constrained model (origin or destination based)

- Both production and attraction are known
  - Doubly constrained model
Direct demand model

Basic gravity model:

\[ T_{ij} = \rho Q_i X_j F_{ij} \]

\[ \sum_i \sum_j T_{ij} = T \]

- \( T_{ij} \) = # trips from zone \( i \) to zone \( j \)
- \( \rho \) = measure of average trip intensity
- \( Q_i \) = production potential of zone \( i \)
- \( X_j \) = attraction potential of zone \( j \)
- \( F_{ij} \) = accessibility of \( j \) from \( i \)

The production potential and attraction potential have to be derived from population, area, jobs (see also OD-matrix estimation).

The production numbers and attraction numbers are unknown.

\[ \rho = \frac{T}{\sum_i \sum_j Q_i \cdot X_j \cdot F_{ij}} \]
Singly constrained model

Basic gravity model:

\[ T_{ij} = \rho Q_i X_j F_{ij} \]

- \( T_{ij} \) = # trips from zone \( i \) to zone \( j \)
- \( \rho \) = measure of average trip intensity
- \( Q_i \) = production potential of zone \( j \)
- \( X_j \) = attraction potential of zone \( i \)
- \( F_{ij} \) = accessibility of \( j \) from \( i \)

If the trip productions \( P_i \) are known:

\[ \sum_j T_{ij} = P_i \]

If the trip attractions \( A_j \) are known:

\[ \sum_i T_{ij} = A_j \]
Singly constrained: origin

\[
\begin{align*}
T_{ij} &= \rho Q_i X_j F_{ij} \\
\sum_j T_{ij} &= P_i
\end{align*}
\]

\[
\sum_j T_{ij} = \sum_j (\rho Q_i X_j F_{ij}) = \rho Q_i \sum_j (X_j F_{ij}) = P_i
\]

\[
\Rightarrow \quad Q_i = \frac{P_i}{\rho \sum_j X_j F_{ij}}
\]

\[
\Rightarrow \quad T_{ij} = \rho \frac{P_i}{\rho \sum_j X_j F_{ij}} X_j F_{ij} = a_i P_i X_j F_{ij}
\]

(\(a_i = \) balancing factor)

Singly constrained origin based model: \(T_{ij} = a_i P_i X_j F_{ij}\)
Singly constrained: destination

\[
\begin{align*}
T_{ij} &= \rho Q_i X_j F_{ij} \\
\sum_i T_{ij} &= A_j
\end{align*}
\]

\[
\sum_i T_{ij} = \sum_i (\rho Q_i X_j F_{ij}) = \rho X_j \sum_i (Q_i F_{ij}) = A_j
\]

\[
\Rightarrow X_j = \frac{A_j}{\rho \sum_i Q_i F_{ij}}
\]

\[
\Rightarrow T_{ij} = \rho Q_i \frac{A_j}{\rho \sum_i Q_i F_{ij}} F_{ij} = b_j Q_i A_j F_{ij}
\]

(\(b_j = \text{balancing factor}\))

Singly constrained destination based model: \(T_{ij} = b_j Q_i A_j F_{ij}\)
Doubly constrained model

Basic gravity model:

\[ T_{ij} = \rho Q_i X_j F_{ij} \]

- \( T_{ij} \) = # trips from zone \( i \) to zone \( j \)
- \( \rho \) = measure of average trip intensity
- \( Q_i \) = production potential of zone \( j \)
- \( X_j \) = attraction potential of zone \( i \)
- \( F_{ij} \) = accessibility of \( j \) from \( i \)

Trip productions \( P_i \) and trip attractions \( A_j \) are known:

\[ \sum_j T_{ij} = P_i \quad \text{and} \quad \sum_i T_{ij} = A_j \]
Doubly constrained model

\[
\begin{aligned}
T_{ij} &= \rho Q_i X_j F_{ij} \\
\sum_i T_{ij} &= A_j \\
\sum_j T_{ij} &= P_i
\end{aligned}
\]

\[
\begin{aligned}
\sum_j T_{ij} &= \sum_j \left( \rho Q_i X_j F_{ij} \right) = \rho Q_i \sum_j \left( X_j F_{ij} \right) = P_i \\
\sum_i T_{ij} &= \sum_i \left( \rho Q_i X_j F_{ij} \right) = \rho X_j \sum_i \left( Q_i F_{ij} \right) = A_j
\end{aligned}
\]

\[
Q_i = \frac{P_i}{\rho \sum_j \left( X_j F_{ij} \right)} \quad \text{and} \quad X_j = \frac{A_j}{\rho \sum_i \left( Q_i F_{ij} \right)}
\]
Doubly constrained model

\[ T_{ij} = \rho \frac{P_i}{\rho} \sum_j X_{ij} F_{ij} \cdot \frac{A_j}{\rho} \sum_i Q_{ij} F_{ij} = a_i b_j P_i A_j F_{ij} \]

\( a_i \) = balancing factor
\( b_j \) = balancing factor
2.1b

Method 1b: Entropy based
Maximising entropy given constraints

Analogue to the thermodynamic concept of entropy as maximum disorder, the entropy-maximizing procedure seeks the most likely configuration of elements within a constrained situation.

The objective can be formulated as:

$$\text{Max } w(T_{ij}) = \frac{T!}{\prod_{ij} T_{ij}!}$$

$$\sum_j T_{ij} = P_i$$

$$\sum_i T_{ij} = A_j$$

$$\sum_i \sum_j T_{ij} \cdot c_{ij} = C$$
Illustration entropy principle

• How many ways can you distribute 4 people?

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• Let’s assume we use a coin to decide where a person will go to: thus flip a coin 4 times

• In total there are 16 sequences leading to 5 options:
  • 4H,0T (1)  3H,1T (4), 2H,2T (6), 1H,3T (4), 0H,4T (1)

• Weight of each option is determined by

\[
\frac{T!}{\prod_{ij} T_{ij}!}
\]
Derivation (1/2)

Lagrangian of maximisation objective:

$$\frac{T!}{\prod_{ij} T_{ij}!} + \sum_i \lambda_i \left( P_i - \sum_j T_{ij} \right) + \sum_j \lambda_j \left( A_j - \sum_i T_{ij} \right) + \beta \left( C - \sum_i \sum_j T_{ij} \cdot c_{ij} \right)$$

Using as approximation

$$\ln(N!) \approx N \cdot \ln(N) - N \Rightarrow \frac{\partial \ln(N!)}{\partial N} \approx \ln(N)$$

Setting derivatives equal to zero and solving the equation

$$\frac{\partial T}{\partial T_{ij}} = -\ln(T_{ij}) - \lambda_i - \lambda_j - \beta \cdot c_{ij} = 0$$

$$\Rightarrow T_{ij} = e^{-\lambda_i - \lambda_j - \beta \cdot c_{ij}} = e^{-\lambda_i} \cdot e^{-\lambda_j} \cdot e^{-\beta \cdot c_{ij}}$$

Note that the other derivatives lead to the original constraints
Derivation (2/2)

Substitute result in constraints

\[ \sum_j T_{ij} = \sum_j e^{-\lambda_i - \lambda_j - \beta c_{ij}} = e^{-\lambda_i} \sum_j e^{-\lambda_j - \beta c_{ij}} = P_i \]

\[ \Rightarrow \frac{e^{-\lambda_i}}{P_i} = \frac{1}{\sum_j e^{-\lambda_j - \beta c_{ij}}} = a_i \]

Similar for destinations, and substitute in \( T_{ij} \)

\[ T_{ij} = e^{-\lambda_i} \cdot e^{-\lambda_j} \cdot e^{-\beta c_{ij}} = a_i \cdot P_i \cdot b_j \cdot A_j \cdot e^{-\beta c_{ij}} \]

Which is equivalent to the doubly constrained model

Thus different routes lead to similar formulation
2.1c

Method 1: Distribution functions and example
Distribution or deterrence functions

\[ F_{ij} = f(c_{ij}) \]

\( c_{ij} = \) travel cost from zone \( i \) to zone \( j \)

distribution function
describes the relative willingness to make a trip as a function of the travel costs.
Requirements for distribution functions

- Decreasing with travel costs
- Integral should be finite

\[
\frac{F(a \cdot c_{ij})}{F(c_{ij})} \text{ depends on value of } c_{ij}
\]

- Fixed changes should have a diminishing relative impact:

\[
\frac{F(c_{ij} + \Delta c)}{F(c_{ij})} > \frac{F(c_{ij} + A + \Delta c)}{F(c_{ij} + A)}
\]
Distribution functions

Power function: \( f(c_{ij}) = c_{ij}^{-\alpha} \)

Exponential function: \( f(c_{ij}) = \alpha \cdot \exp(-\beta c_{ij}) \)

Top-exponential function: \( f(c_{ij}) = \alpha c_{ij}^\beta \cdot \exp(-\gamma c_{ij}) \)

Lognormal function: \( f(c_{ij}) = \alpha \cdot \exp(-\beta \cdot \ln^2(c_{ij} + 1)) \)

Top-lognormal function: \( f(c_{ij}) = \alpha c_{ij}^\beta \cdot \exp(-\gamma \cdot \ln^2(c_{ij} + 1)) \)

Log-logistic function: \( f(c_{ij}) = \alpha/(1 + \exp(\beta + \gamma \cdot \log(c_{ij}))) \)
Example doubly constrained model

![Diagram of a network with nodes 1, 2, and 3, connected by paths with labeled costs.]

<table>
<thead>
<tr>
<th>$P_i$</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

| $A_j$ | 220 | 165 | 165 |

Trip balancing
Example doubly constrained model

Trip balancing

Travel costs $c_{ij}$

<table>
<thead>
<tr>
<th></th>
<th>$P_i$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>250</td>
</tr>
<tr>
<td>$A_j$</td>
<td>220</td>
<td>165</td>
<td>165</td>
<td></td>
</tr>
</tbody>
</table>
Example doubly constrained model

Trip balancing

Travel costs $c_{ij}$

Accessibilities $F_{ij}$

$F_{ij} = f(c_{ij})$

$= 5 \cdot \exp(-0.5 \cdot c_{ij})$
Example doubly constrained model

Trip balancing
Travel costs $c_{ij}$
Accessibilities $F_{ij}$
Balancing factors $a_i, b_j$

<table>
<thead>
<tr>
<th></th>
<th>3.0</th>
<th>1.1</th>
<th>1.1</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_i$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x\frac{100}{5.2}$</td>
<td>3.0</td>
<td>1.1</td>
<td>1.1</td>
<td>100</td>
</tr>
<tr>
<td>$x\frac{200}{7.1}$</td>
<td>1.1</td>
<td>3.0</td>
<td>3.0</td>
<td>200</td>
</tr>
<tr>
<td>$x\frac{250}{6.6}$</td>
<td>1.8</td>
<td>1.8</td>
<td>3.0</td>
<td>250</td>
</tr>
<tr>
<td>$A_j$</td>
<td>220</td>
<td>165</td>
<td>165</td>
<td></td>
</tr>
</tbody>
</table>
Example doubly constrained model

Trip balancing

Travel costs $c_{ij}$

Accessibilities $F_{ij}$

Balancing factors $a_i, b_j$

\[
\begin{array}{cccc}
57.6 & 21.2 & 21.2 & 100 \\
31.0 & 84.5 & 84.5 & 200 \\
68.2 & 68.2 & 113.6 & 250 \\
\end{array}
\]

\[
\begin{array}{cccc}
220 & 165 & 165 \\
220 & 165 & 165 \\
\end{array}
\]

\[
\frac{173.9}{219.3}
\]
Trip distribution using the gravity model

Trip balancing

Travel costs $c_{ij}$

Accessibilities $F_{ij}$

Balancing factors $a_i, b_j$

<table>
<thead>
<tr>
<th></th>
<th>$P_i$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>80.8</td>
<td>20.1</td>
<td>15.9</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>43.5</td>
<td>80.2</td>
<td>63.6</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>95.7</td>
<td>64.7</td>
<td>85.5</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$A_j$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>165</td>
<td>165</td>
<td></td>
</tr>
</tbody>
</table>
Example doubly constrained model

Trip balancing
Travel costs $c_{ij}$
Accessibilities $F_{ij}$
Balancing factors $a_i, b_j$
Repeat until there are no changes:
$\Rightarrow$ OD matrix
2.2

Method 2: Extrapolation or growth
Growth factor models

- Base matrix
  - Trip observations
  - Outcomes of gravity model
  - Traffic counts

Growth factors
  - Outcomes future trip generation model
  - Outcomes future gravity model

Predicted matrix
Growth factor models

Advantages
• Network specific peculiarities can be captured by observations
• A base matrix is more understandable and verifiable than a model

Disadvantages
• New residential zones are difficult to capture
• Historical patterns may change over time
Growth factor models

\[ T_{ij}^0 \rightarrow g \rightarrow T_{ij} = gT_{ij}^0 \]

- Base matrix cell: \( T_{ij}^0 \)
- Growth factor: \( g \)
- Predicted matrix cell: \( T_{ij} \)

\[ T_{ij} = gT_{ij}^0 \]

Network independent, general factor

\[ T_{ij} = g_i T_{ij}^0 \quad \text{or} \quad T_{ij} = g_j T_{ij}^0 \]

Network independent, origin or destination specific factor

\[ T_{ij} = g_i g_j T_{ij}^0 \]

Network independent, factors for origins and destinations

\[ T_{ij} = g_{ij} T_{ij}^0 \]

Network dependent, OD-specific factors
Common application

- Given expected spatial development
  - Future production (departures)
  - Future attraction (arrivals)

- Fill in new areas by copying columns and rows of nearby zones from the base year matrix

- Adapt this base year matrix using appropriate factors $a_i$ and $b_j$
  Iteration process slides 33-35
2.3

*Method 3: Choice modelling*
Discrete choice model

\[ T_{ij} = P_i \frac{\exp(\mu V_j)}{\sum_k \exp(\mu V_k)}, \quad V_j = \beta_1 X_j - \beta_2 c_{ij} \]

- \( T_{ij} \): number of trips from \( i \) to \( j \)
- \( \beta_1, \beta_2 \): parameters
- \( \mu \): scaling parameter
- \( P_i \): trip production at zone \( i \)
- \( X_j \): trip attraction potential at zone \( j \)
- \( c_{ij} \): travel cost from zone \( i \) to zone \( j \)
Explanatory variables?

- Inhabitants
- Households
- Jobs
- Retail jobs
- Students
- Densities
- Location types
- Etc.

- Minus travel costs

More suited for trips or for tours?
Derivation of the gravity model (reprise)

Observed utility for activities in zone $i$ and zone $j$:

$$V_{ij} = N_j - N_i - \beta_2 \cdot c_{ij}$$

Subjective utility:

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

Number of people traveling from $i$ to $j$:

$$p_{ij} \cdot T = \frac{e^{\mu V_{ij}} \cdot T}{\sum_{rs} e^{\mu V_{rs}}} = \frac{T \cdot \exp(-\mu N_i) \cdot \exp(\mu N_j) \cdot \exp(-\mu \beta_2 c_{ij})}{\sum \exp(\mu V_{rs})}$$

$$T_{ij} = \rho Q_i X_j F_{ij}$$
3.

Special issues
Special issues

- Distribution function and trip length distribution
- Intrazonal trips
- External zones: through traffic
- All trips or single mode?
Distribution function and trip length distribution

- Similar or different?

- Simply put: distribution function is input and trip length distribution is output!
Intrazonal trips

- What’s the problem?

- Intrazonal travel costs?

- Rule of thumb: \( \frac{1}{3} \) (or \( \frac{1}{2} \)?) of lowest cost to neighbouring zone
  - True for public transport?

- Alternative: Trip generation for intrazonal only
  - How?
External zones

- Two possible issues

- Size issue
  - Very large zones => high values for production and attraction
    => intrazonal trips? => small errors lead to large differences

- Cordon models
  - Through traffic follows from other source, e.g. license plate survey or other model => through traffic is thus fixed input and should not be modelled using trip distribution models
Approach for cordon model

- Determine production and attraction for internal zones using e.g. regression
- Determine production and attraction for external zones using e.g. counts
- Derive matrix for through traffic (i.e. from cordon zone to cordon zone) from e.g. a regional model
- Subtract through traffic from production and attraction of the external zones
- Apply gravity model with the resulting production and attraction, while making sure that there is no through traffic, e.g. by setting the travel costs between cordon zones equal to $\infty$
- Add matrix for through traffic to the resulting matrix of the gravity model
All trips or a single mode?

• Check the slides

• Which parts consider a single mode?