CIE4801 Transportation and spatial modelling
4-step model reprise, forecasting

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

• 4 Step model: reprise
  • Your opinion on the models we discussed
  • Some comments
    • Aggregate and disaggregate models
    • Feedbacks
    • Brasess paradox
    • Multimodality
    • Nested Logit

• Using the model in a study
  • Matrix adjustment (growth factor)
  • Just running the model
  • Pivot point method

• Model data and your (possible) MSc-thesis
  • What to do if you don’t get what you want
1.

4 Step-model: reprise
Set-up course

- Theory
  - Reader
  - Lectures

- Practice
  - Lectures by experts from the field

- Experience
  - Practical
  - Exam questions
    - Exercises
    - Delft case

First: Transportation modelling
Second: Spatial modelling
Alternative scheme of 4-step model
How good is it?

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.
Requirements of lecture 1

- Proper representation of the transportation system or capture the key mechanisms of the system
- Theoretically underpinned
- Empirically validated
- As simple as possible, but not simpler
# Overview of approaches

<table>
<thead>
<tr>
<th></th>
<th>Descriptive</th>
<th>Choice behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trip generation</td>
<td>Regression</td>
<td>Cross classification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stop &amp; Go</td>
</tr>
<tr>
<td>Distribution</td>
<td>Analogy</td>
<td>Growth factor model</td>
</tr>
<tr>
<td></td>
<td>Gravity model</td>
<td>Destination choice</td>
</tr>
<tr>
<td></td>
<td>Entropy model</td>
<td></td>
</tr>
<tr>
<td>Modal split</td>
<td>VF-ratio</td>
<td>Mode choice</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>Gravity model</td>
<td></td>
</tr>
<tr>
<td>distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>modal split</td>
<td>$F_{ij} = \sum_v F_{ijv}$</td>
<td></td>
</tr>
<tr>
<td>Time of day</td>
<td>Fixed shares</td>
<td>TOD choice</td>
</tr>
<tr>
<td>Assignment</td>
<td>AON</td>
<td>DUE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple routing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SUE</td>
</tr>
</tbody>
</table>
Possible segmentations?

<table>
<thead>
<tr>
<th>Trip generation</th>
<th>Trip purpose</th>
<th>Car ownership</th>
<th>Person characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>Trip purpose</td>
<td>Mode</td>
<td></td>
</tr>
<tr>
<td>Modal split</td>
<td>Car ownership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simultaneous</td>
<td>Mode?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>distribution model split</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day</td>
<td>Trip purpose?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assignment</td>
<td>????</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overview special issues

<table>
<thead>
<tr>
<th>Zones</th>
<th>Cordon model</th>
<th>Study, influence and external area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network</td>
<td>Connectors</td>
<td></td>
</tr>
<tr>
<td>Trip generation</td>
<td>External zones</td>
<td>Balancing</td>
</tr>
<tr>
<td>Distribution</td>
<td>Intrazonal costs</td>
<td>Through traffic</td>
</tr>
<tr>
<td>Modal split</td>
<td>Car ownership</td>
<td>Car occupancy</td>
</tr>
<tr>
<td>Simultaneous distribution model split</td>
<td>See above</td>
<td></td>
</tr>
<tr>
<td>Time of day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assignment</td>
<td>Trucks</td>
<td></td>
</tr>
</tbody>
</table>
2.

Some comments
Aggregate/ disaggregate models, feedbacks, multimodality and nested logit
Agregate and disaggregate models

Main concept

Single group
=> Average traveller

Individual

Common interpretation

Descriptive models

Choice models

Practice

Descriptive models for sets of groups

Choice models for groups of individuals
Feedbacks

- 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.

DUE, SUE MSA, Frank-Wolfe

Naive Maximum iterations

CIE4801: Reprise and forecasting
Braess paradox in traffic assignment

\[ t_1(q_1) = \frac{q_1}{100} \]

\[ t_2(q_2) = \frac{q_2}{100} \]

\[ T_{AB} = 1000 \]

DUE:

\[ f_1 = 500 \]
\[ f_2 = 500 \]
\[ t_1^{AB} = t_2^{AB} = 20 \]

DUE with new (green) connection:

\[ f_1 = 250 \]
\[ f_2 = 250 \]
\[ f_3 = 500 \]
\[ t_1^{AB} = t_2^{AB} = t_3^{AB} = 22.5 \]
Multimodality?

- Approach so far: clear split between car, public transport (and slow modes)
- Special case: separate modes for train and BTM
- 80% of train travellers use other modes for access and/or egress
  - Bus/tram/metro, but also bike and car
- So where do we find the cyclist to the station?
Multimodal trips

- **Car network**
- **Public transport network**
- **Pedestrian network**

Origin to destination.
Simultaneous route/mode choice

- car
- train
- bus
- bike

super network
Alternative model structures

- Production/attraction
- Distribution
- Modal split
- Assignment

- Production/attraction
- Distribution
- Modal split
- Assignment

- Production/attraction
- Distribution
- Modal split
- Assignment
Nested logit: Classic example

- Red and blue bus problem
Coping with correlations: Nesting

- Nesting accounts for (unobserved) similarities within nests: mix of correlation, simultaneousness and hierarchy
  - It does not necessarily imply a sequential order of choices!

- Special application/interpretation: Conditional choice:
  - Choice for alternative given choice for nest
  - Lower level choice options are part of higher level utility
Typical example

\[ P_i = \frac{e^{V_i}}{\sum_j e^{V_j}} \]

\[ P_i = \frac{e^{V_i/\lambda_k} \left( \sum_{j \in B_k} e^{V_j/\lambda_k} \right)^{\lambda_k - 1}}{\sum_{l=1}^K \left( \sum_{j \in B_l} e^{V_j/\lambda_l} \right)^{\lambda_l - 1}} \]
Decomposition in two logits

Split utility in two parts:

- variables describing attributes for nests (aggregate level): \( W_k \)
- variables describing attributes within nest: \( Y_j \)

\[
U_j = W_k + Y_j + \varepsilon_j \quad j \in B_k
\]

Probability alternative is product of probability of alternative within nest and probability of nest

\[
P_i = P_{i|B_k} P_{B_k}
\]
Decomposition in two logits
Resulting formulas

\[ P_{B_k} = \frac{e^{W_k + \lambda_k I_k}}{\sum_{l=1}^{K} e^{W_l + \lambda_l I_l}} \]

\[ P_{i|B_k} = \frac{e^{Y_i / \lambda_k}}{\sum_{j \in B_k} e^{Y_j / \lambda_k}} \]

\[ I_k = \ln \sum_{j \in B_k} e^{Y_j / \lambda_k} \]
Logsum:  
“translation” of aggregate quality

\[ v_i = \frac{1}{\mu} \ln \left( \sum_{n} \exp \left( \mu \cdot V_i \right) \right) \]

\[ = \bar{V} + \frac{1}{\mu} \ln \left( \frac{\sum_{n} \exp \left( \mu \cdot (V_i - \bar{V}) \right)}{n} \right) + \frac{1}{\mu} \ln (n) \]

\[ v_i = \text{Systematic utility according to logsum} \]
\[ \bar{V} = \text{Average systematic utility} \]
\[ V_i = \text{Systematic utility alternative i} \]
\[ \mu = \text{Scale parameter} \]
\[ n = \text{Number of alternatives} \]
Aggregated travel time for 2 routes
Travel time route 1 is 40 minutes

![Travel time graph]

- **TT Mean**
- **TT Min**
- **TT Logsum**
Nested logit and hierarchical choices

- Logsum is component in utility of higher level choice model with parameter $\mu^*$

\[ \mu^* \cdot \ln \left( \frac{1}{\mu} \ln \left( \sum_n \exp(\mu \cdot V_i) \right) \right) \]

- $0 < \mu^*/\mu \leq 1$: not relevant to fully independent (NB $\mu^*/\mu \approx \lambda$)

- Limitation: alternative can only be allocated to a single nest
Examples

- Logsum over routes in mode choice
- Logsum over modes in destination choice
- Logsum over activities in activity choice
  - Activity based modelling
- Dutch National Model considers nesting when modelling destination and mode choice (and tour generation)
Four stage model and logsums

Trip generation

Destination A
  - Mode 1
    - Route I
    - Route II
  - Mode 2
    - Route I
    - Route II

Destination B
  - Mode 1
  - Mode 2

Destination C
  - Mode 1
  - Mode 2

CIE4801: Reprise and forecasting
3. Applying a model in practice
Applying a model in practice

- Base year model
  - Justification of your model

- However, model is a simplification of reality. How to deal with that when making a forecast?

- 3 options
  - Rely on base year matrix
  - Rely on model
  - Combination of both

- In all cases: proper interpretation of the results
Growth factor models: reprise

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

- **base matrix cell**
- **growth factor**
- **predicted matrix cell**

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

**Network independent, general factor**

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

**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**
Rely on the base year matrix

- Make sure that the base year is of high quality
  - Good fit with counts using a high quality a priori matrix

- Use a growth factor method for the forecast
  - Especially changes in production and attraction \( (g_i, g_j) \)

- Limitations
  - Network effects on distribution and modal split are not included
  - Mostly suitable for short term forecasts only
Rely on the model

- Make sure that the model is of high quality
  - Proper functions and parameters
  - Decent match with counts etc.

- Run the model using the input for the forecast year

- Check the model for those results where the base year model had some limitations
Combination of both worlds

• Key philosophy:
  • Model is the best representation of the mechanisms in the transport system, i.e. of the sensitivity for measures you’re interested in
  • Base year matrix is best representation of the spatial pattern

• Thus use the model to determine the relative changes between base year and forecast year (synthetic model or shadow model)

• Next, use these relative changes to update the base year matrix (pivot) \((g_{ij})\)
Pivot model approach

Base year matrix

Input base year model

Base year model matrix

Input forecast year model

Forecast year model matrix

Growth factor matrix \((g_{ij})\)

Forecast year matrix

Calibrated model
Points of attention for the pivot method

- Relationship between base year matrix and base year model matrix?

- Growth factor method extrapolates “errors” in the base year matrix
  => Demand patterns should be consistent

- How to account for changes in spatial structure (empty cell problem)
  - There’s nothing to multiply with

Substitute cell value with results synthetic OD-matrix
Or,
First, use matrix manipulation to account for changes in spatial structure (see lecture 3) and then use pivot to account for network and policy changes only
In all cases

• Take care of the relevant accuracy
  • Within the model there are partial persons and partial cars!
  • Exact values tend to become a fact

• Use your common sense!
  • What’s the logic?
  • Do you believe the changes and their magnitudes?
  • Can you explain it in laymen’s terms as well?
4.

Using a model in a MSc-thesis project
For many MSc studies you would like to use a model

- Option 1: You can’t get any modelling data
- Option 2: There’s a model and you can get networks and OD-matrices
- Option 3: You’ve got access to the full modelling framework
No modelling data at all

- Use other data to get an idea of the main transport characteristics
  - Counts
  - Household surveys, e.g. OVG/MON/OViN
  - Reports on modelling studies in the same area

- Use engineering judgement
  - For typical OD-pairs determine the level of service (skims of travel times etc) and use e.g. a logit model to estimate mode choice or route choice (the parameters are your choice)
You’ve got networks and matrices

- You can do matrix manipulations and assignments
  - Usually you’re interested in a part of the model
    => adapt the zoning system

- What’s missing is e.g. a mode choice model
- 2 options
  - Determine level of service using the networks (skims) and apply a (self-made) version of the Poisson-model
  - Determine level of service and estimate a mode choice model (e.g. a Logit model having mode specific constants and parameters for the level of service of the modes)

- In both cases select a set of relevant OD-pairs that are consistent in size
You’ve got the full framework

- Don’t get lost in all the data and all modelling features
- You might want to adapt the zoning system to your study goal
5.

Coming next: Guest lectures
Next lectures

• Lectures from practice in The Netherlands
  • Dutch National Model System: Gerard de Jong (Significance, ITS Leeds)
  • Base year matrix estimation (Regional Model): Peter Mijjer (4Cast)
  • Regional Model application: Jan van der Waard (KIM)
  • Urban transport modelling: Bastiaan Possel (Goudappel Coffeng)

• Goal of these presentations:
  • Lectures focus on basic theory, practical on a simple case
  • These lectures focus on models in practice and their context

• Your task:
  • Determine similarities and differences between my story and theirs