CIE4801 Transportation and spatial modelling
Trip distribution

Rob van Nes, Transport & Planning

17/4/13

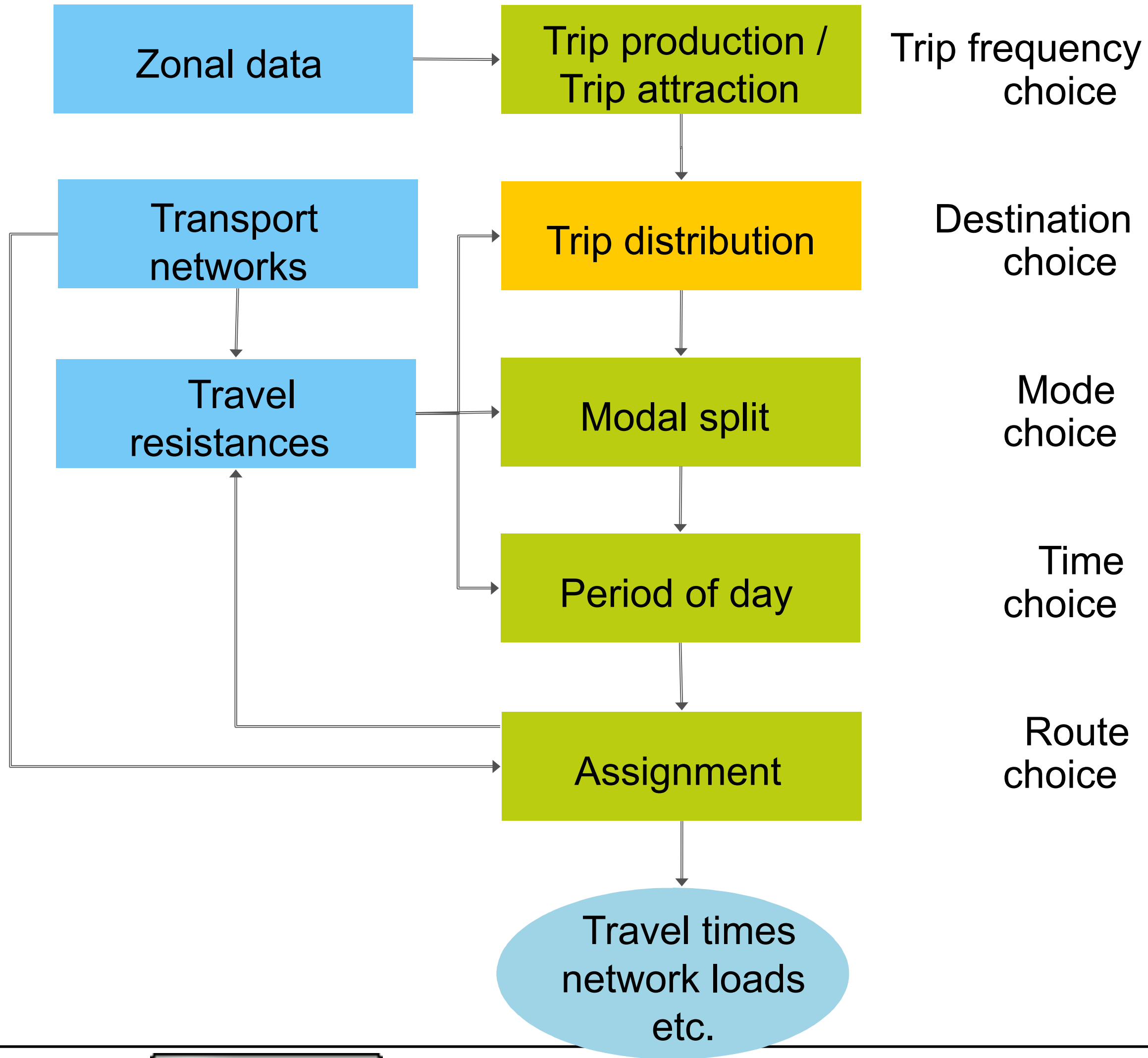
Content

- 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

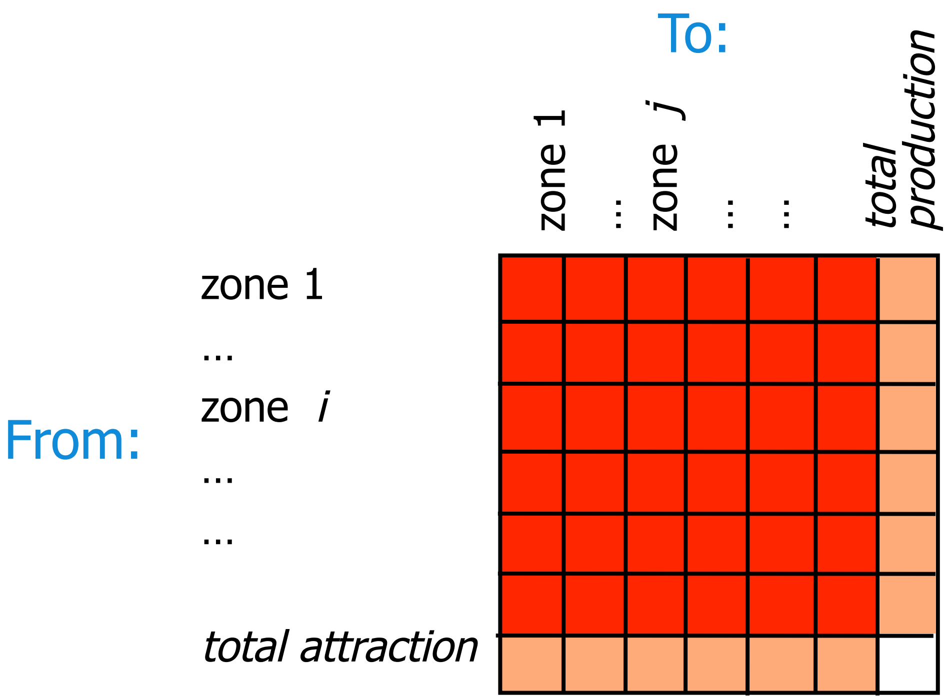


What do we want to know?

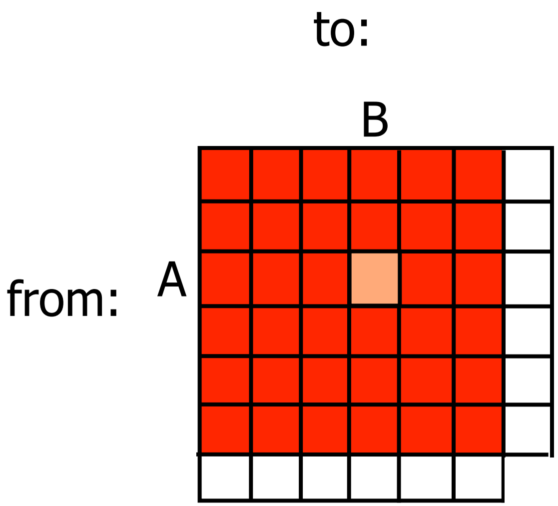
Introduction to trip distribution

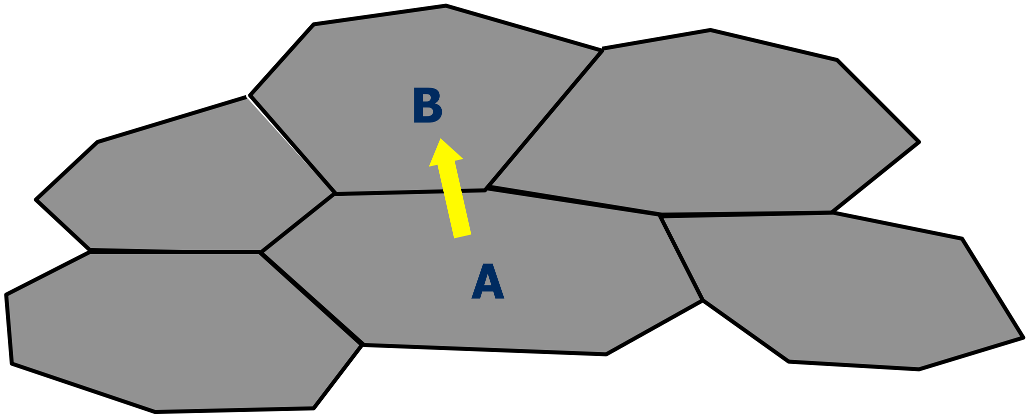
Given: Productions and attractions for each zone (i.e. departures and arrivals)

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



Introduction to trip distribution





Trip distribution: 3 methods

- Analogy based
- Extrapolation or growth
- Choice modelling

2.1

Method 1: Analogy

The gravity model

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

m_i

d_{ij}

m_j

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*

ρ = 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*

ρ = 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*

ρ = 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

$$\begin{cases} T_{ij} = \rho Q_i X_j F_{ij} \\ \sum_j T_{ij} = P_i \end{cases}$$

$$\sum_j T_{ij} = \sum_j \left(\rho Q_i X_j F_{ij} \right) = \rho Q_i \sum_j \left(X_j F_{ij} \right) = P_i$$

$$\Rightarrow \qquad Q_i = \frac{P_i}{\rho \sum_j X_j F_{ij}}$$

$$\Rightarrow \qquad T_{ij} = \rho \frac{P_i}{\rho \sum_j X_j F_{ij}} X_j F_{ij} = a_i P_i X_j F_{ij}$$

(a_i = balancing factor)

Singly constrained origin based model: $T_{ij} = a_i P_i X_j F_{ij}$

Singly constrained: destination

$$\begin{cases} T_{ij} = \rho Q_i X_j F_{ij} \\ \sum_i T_{ij} = A_j \end{cases}$$

$$\sum_i T_{ij} = \sum_i \left(\rho Q_i X_j F_{ij} \right) = \rho X_j \sum_i \left(Q_i F_{ij} \right) = A_j$$

$$\Rightarrow \qquad X_j = \frac{A_j}{\rho \sum_i Q_i F_{ij}}$$

$$\Rightarrow \qquad T_{ij} = \rho Q_i \frac{A_j}{\rho \sum_i Q_i F_{ij}} F_{ij} = b_j Q_i A_j F_{ij}$$

(b_j = balancing factor)

Singly constrained destination based model: $T_{ij} = b_j Q_i A_j F_{ij}$

Doubly constrained model

<u>Basic gravity model:</u>

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

T_{ij} = # trips from zone *i* to zone *j*

ρ = 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

$$\begin{cases} T_{ij} = \rho Q_i X_j F_{ij} \\ \sum_i T_{ij} = A_j \\ \sum_j T_{ij} = P_i \end{cases}$$

$$\sum_j T_{ij} = \sum_j \left(\rho Q_i X_j F_{ij} \right) = \rho Q_i \sum_j \left(X_j F_{ij} \right) = P_i$$

$$\sum_i T_{ij} = \sum_i \left(\rho Q_i X_j F_{ij} \right) = \rho X_j \sum_i \left(Q_i F_{ij} \right) = A_j$$

$$\Rightarrow \qquad Q_i = \frac{P_i}{\rho \sum_j \left(X_j F_{ij} \right)} \qquad \text{and} \qquad X_j = \frac{A_j}{\rho \sum_i \left(Q_i F_{ij} \right)}$$

Doubly constrained model

$$\Rightarrow \quad T_{ij} = \rho \frac{P_i}{\rho \sum_j X_j F_{ij}} \cdot \frac{A_j}{\rho \sum_i Q_i F_{ij}} F_{ij} = a_i b_j P_i A_j F_{ij}$$

a_i = balancing factor

b_j = balancing factor

Doubly constrained model: $T_{ij} = a_i b_j P_i A_j F_{ij}$

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:

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

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

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

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

Illustration entropy principle

- How many ways can you distribute 4 people?

	H	T
C		

- 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:

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

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 orginal constraints

Derivation (2/2)

Substitute result in constraints

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

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

Similar for destinations, and substitute in T_{ij}

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

Which is equivalent to the doubly constrained model
Thus different routes lead to similar formulation

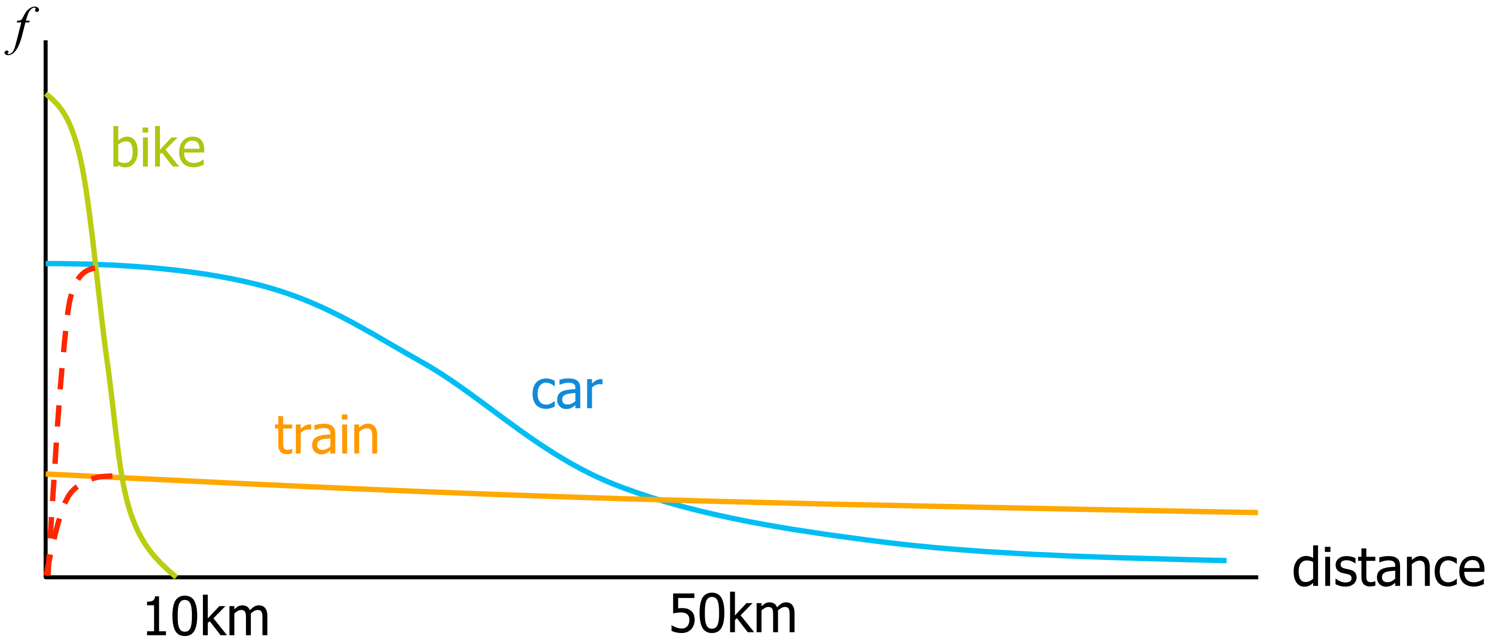
2.1c

Method 1: Distribution functions and example

Distribution or deterrence functions

$F_{ij} = \underbrace{f(c_{ij})}$ $\qquad c_{ij} =$ 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
- Fraction $\frac{F(a \cdot c_{ij})}{F(c_{ij})}$ 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: $f(c_{ij}) = c_{ij}^{-\alpha}$

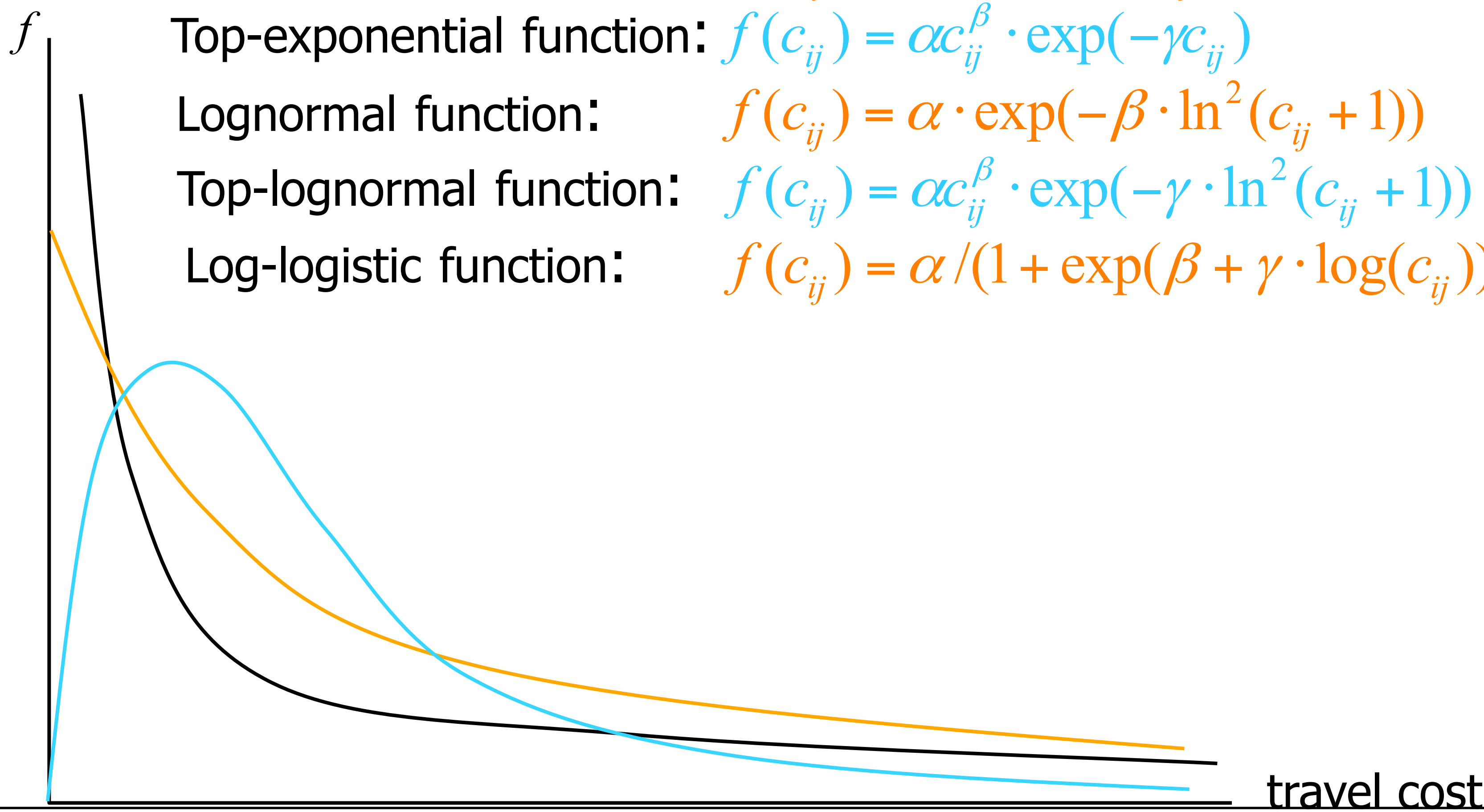
Exponential function: $f(c_{ij}) = \alpha \cdot \exp(-\beta c_{ij})$

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

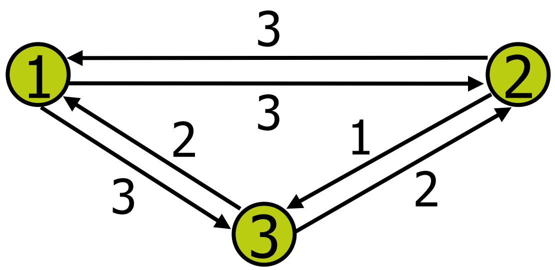
Lognormal function: $f(c_{ij}) = \alpha \cdot \exp(-\beta \cdot \ln^2(c_{ij} + 1))$

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

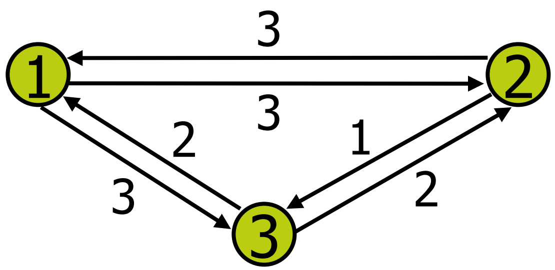


Trip balancing

			P_i
			100
			200
			250
220	165	165	

A_j

Example doubly constrained model

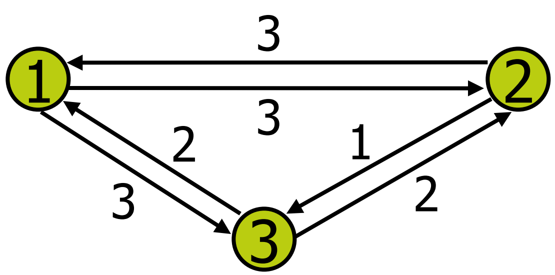


Trip balancing

Travel costs c_{ij}

				P_i
	1	3	3	100
	3	1	1	200
	2	2	1	250
A_j	220	165	165	

Example doubly constrained model



Trip balancing

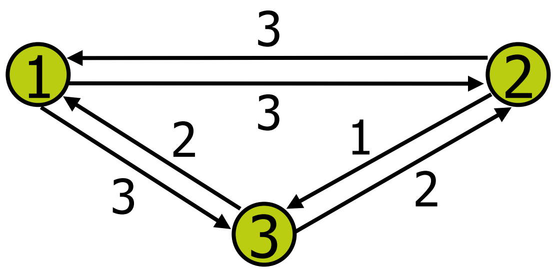
Travel costs c_{ij}

Accessibilities F_{ij}

				P_i
	3.0	1.1	1.1	100
	1.1	3.0	3.0	200
	1.8	1.8	3.0	250
A_j	220	165	165	

$$F_{ij} = f(c_{ij})$$
$$= 5 \cdot \exp(-0.5 \cdot c_{ij})$$

Example doubly constrained model



				P_i
$\times\frac{100}{5.2}$	3.0	1.1	1.1	100
$\times\frac{200}{7.1}$	1.1	3.0	3.0	200
$\times\frac{250}{6.6}$	1.8	1.8	3.0	250
A_j	220	165	165	

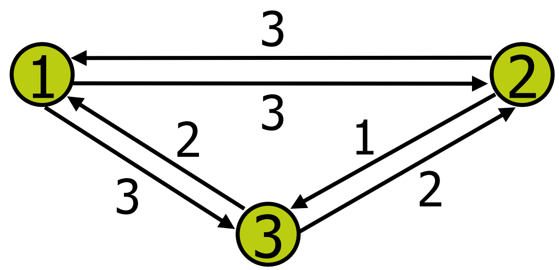
Trip balancing

Travel costs c_{ij}

Accessibilities F_{ij}

Balancing factors a_i, b_j

Example doubly constrained model



Trip balancing

Travel costs c_{ij}

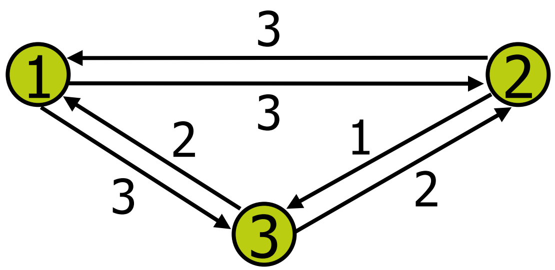
Accessibilities F_{ij}

Balancing factors a_i, b_j

				P_i
	57.6	21.2	21.2	100
	31.0	84.5	84.5	200
	68.2	68.2	113.6	250
A_j	220	165	165	

$$\times\frac{220}{156.8} \quad \times\frac{165}{173.9} \quad \times\frac{165}{219.3}$$

Trip distribution using the gravity model



				P_i
	80.8	20.1	15.9	100
	43.5	80.2	63.6	200
	95.7	64.7	85.5	250
A_j	220	165	165	

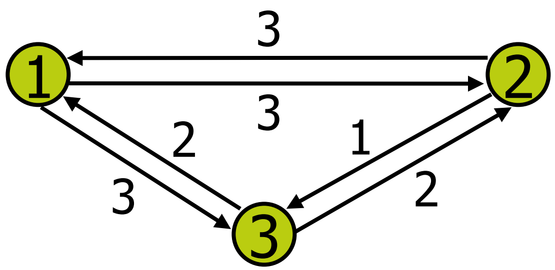
Trip balancing

Travel costs c_{ij}

Accessibilities F_{ij}

Balancing factors a_i, b_j

Example doubly constrained model



				P_i
	70.5	16.5	13.0	100
	48.7	84.3	67.0	200
	100.8	64.2	85.0	250
A_j	220	165	165	

Trip balancing

Travel costs c_{ij}

Accessibilities F_{ij}

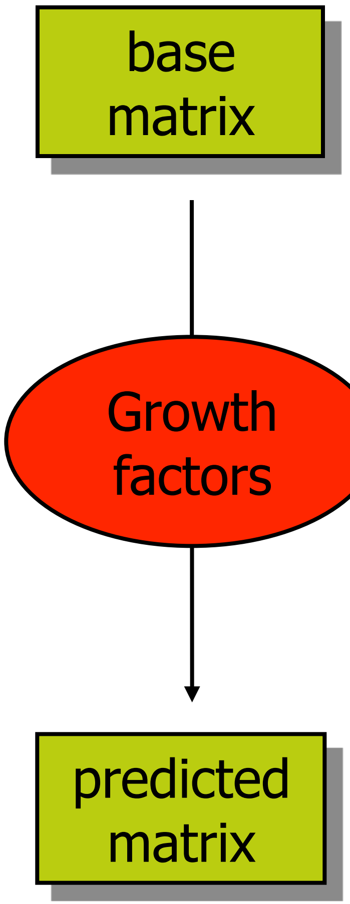
Balancing factors a_i, b_j

Repeat until there are no changes: =>OD matrix

2.2

Method 2: Extrapolation or growth

Growth factor models



- trip observations
- outcomes of gravity model
- traffic counts

- outcomes future trip generation model
- outcomes future gravity model

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

$$T_{ij}^0 \longrightarrow g \longrightarrow T_{ij} = gT_{ij}^0$$

base matrix cell — growth factor — predicted matrix cell

$$T_{ij} = gT_{ij}^0$$

Network independent,
general factor

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

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

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

$T_{ij} =$	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*

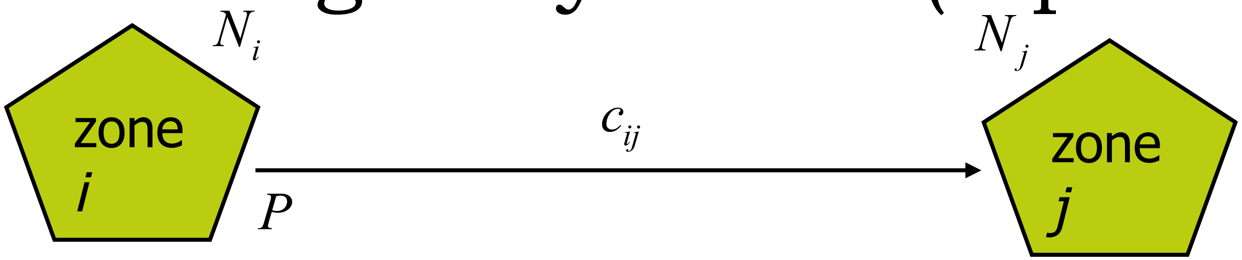
Explanatory variables?

- 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*:

$$p_{ij} \cdot T = \frac{e^{\mu V_{ij}}}{\sum_{rs} e^{\mu V_{rs}}} \cdot T \qquad T_{ij} = \rho Q_i X_j F_{ij}$$

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

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 ∞
- 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?