Transport and Mobility

For a cleaner more efficient future

November 18, 2020

1 Demand

1.1 The Four Step Model

Before you can apply the four step model you need to first:

- Tessellate - discretize the space your going to work with
- Data (Population, demographic data, interaction/flow (migration, commuting, movement data) etc..). UK Census data is considered the 'gold standard'. It gets updated every 10 years.

1.1.1 Trip generation

$ This is about the number of trips originate O and destinate D in a zone.
- O_i is the number of trips 'generated' in zone i.
- D_j is the number of trips 'attracted' to zone j. When zones are iteratively merged, D becomes proportional to population size.
- Trip frequency problem: The probability that a person a certain type will undertake N trips in a week.
- Usually one uses a different four step model for each different type of journey, the lecture (L3) focuses on the travel-to-work journey.
- Trip production factors (the likelihood of a person making a certain trip): Income, Car ownership, Family/household size, land value, residential density etc...
- Trip attraction factors (likelihood of a person ending up, being attracted to somewhere): Employment level at desination zone, accessibility, total amount of floor space for industrial or commercial use, etc...
- Spacial units from UK Census data: LA means Local Authority (e.g. Bristol), Ward means district (or something like that) in local authority (e.g. Clifton, 35 in total in Bristol), MSOA stands for Middle Layer Super Output Area (53 in Bristol, 7201 in England and Wales) [OFTEN BEST ONE TO USE], LSOA stands for Lower Layer Super Output Area (34,753 in England and Wales), OA is the Output Area (181,408 in England and Wales), Post codes (1.7 million in the UK).
- The gradient of the best fitting line for a plot of O against population (for work journey) gives the employment rate.
- The number of origin trips O_i is proportional to the population in zone i and has an upper bound given by the population in zone i.
- The number of destination trips D does not have a natural upper limit.
  - Issues with the data include: Republic of Ireland not included, definitions of boundaries change over time, much can happen in 10 yeas and the census only cover a 12-month period (too general) every 10 years.

1.1.2 Trip distribution

- $ Estimating the origin-destination distributions/ pairings.
- Use goodness of fit: Pearson's Correlation coefficient: $R(X,Y) = cov(X,Y)/(\sigma_X \sigma_Y)$ or Sorensen index (less sensitive to outliers, better suited for non-homogenous data sets)
- Various ways to compute distance → Euclidean probably best for smaller distances (where origin is centre of the earth) and Geodesic for larger distances.
- Gonzalez's paper: `https://www.ole.bris.ac.uk/bbcswebdav/pid-4180487-dt-content-rid-13788613_2/cours EMATM0021_2019_TB-2/nature06958.pdf`: Distribution of trip lengths: $P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} e^{\frac{-\Delta r}{\kappa}}$ (read paper for what variables mean).
- Song's paper `https://www.ole.bris.ac.uk/bbcswebdav/pid-4180487-dt-content-rid-13788614_2/courses/ EMATM0021_2019_TB-2/nphys1760.pdf`: Elapsed time between trips: $P(\Delta t) = \Delta t^{-1-\beta_t} e^{\frac{-\Delta t}{\tau}}$
- Location-visitation frequence (Also Song's paper): A Zipf's law $f_k = k^{-\xi}$
- T_{ij} is the number of tripes from zone i to zone j. There are various different models to estimate T:
  - Gravitation Model: $T_{ij} \propto P_i P_j f(d)$, where P_i and P_j are the populations in zones i and j, and $f(d)$ is often $1/d$ but can also be $1/d^2$ or $e^{-\beta d}$.
    - The model performs better if it is globally, origin, doubly and destination constrained (see lecture 05): `https://www.ole.bris.ac.uk/bbcswebdav/pid-4180487-dt-content-rid-13788616_2/courses/EMATM 2019_TB-2/2087063.pdf`
    - Generalised gravity model: $T_{ij}^{gravity} = A P_i^{\alpha} P_j^{\beta} / d^{\gamma}$ (3 parameters)
  - Radiation Model: $T_{ij}^{radiation} = T_i \frac{P_i P_j}{(P_i + s_{ij})(P_i + P_j + s_{ij}}$ (no parameters)
  - Entropy Maximization,
  - Intervening Opportunities model: $T_{ij}^{io} = (1-\lambda)^{(P_i + s_{ij})} - (1-\lambda)^{(P_i + s_{ij} + P_j)}$ (1 parameter)
- Activity based modelling: Like modelling movement of nurses in a neonatal intensive care unit.
  - Hidden Markov Model got best results with 88% accuracy.

The other two parts of the four step model are: Mode Choice / Modal Split (About modes of transport, e.g. walking, cycling, public transport, etc...) and Traffic Assignment (Route choice) - these are covered in the next section.

2 Assignment

- Supposing demand is fixed, the split of choice between mode (car, bus, train, etc...) and route, is referred to as assignment.
- Factors of assignment include travel time, money, convenience, quality, ethical/ environmental principles. We treat all these factors as a single (virtual) cost. We also assume that the people going from A to B are **rational actors** - they try to minimise their personal cost.
- The **route assignment** is the number of users (amount of traffic) taking a particular route - this is what we want to determine.
- **Wardrop principle** or **User Equilibrium** states that the costs of the used routes will be equal and less than or equal to the costs of the unused routes.
- These are so-called 'congestion games'. The basic idea is that the cost of a choice goes up if more people choose it.
- **Link costs** - cost of using a particular route (link in a network). Given the variable c_i, where i is the link.
- The number of users and their choice is given by x_i where x is the number of users and i is the choice of route. E.g. if $c_1(x_1) = x_1$ then this implies the route is very short (efficient) if the number users is small, but can get easily congested if the number of users increases, whereas if $c_2(x_2) = 1$ for example, it doesn't matter how many users there are, it will always be the same cost (e.g . take the same amount of time, if cost is time for example).
- For a network, we naturally require that $x_1, x_2, ..., x_n \geq 0$ and $x_1 + x_2 + ... + x_n = d$, where d is the demand (i.e.

the total number of users that want to use the network).

- Total system cost $f = x_1c_1(x_1) + x_2c_2(x_2) + ...x_nc_n(x_n)$. This cost can be different when computing it by determining the user equilibrium or the system optimal solution.
- **System Optimal assignment** tries to minimise the total system cost $f = \sum_{i=1}^{n} x_i c_i(x_i)$. The system optimal assignment solution is not necessarily the same as the user equilibrium.
- The **Price of Anarchy** is defined as $PoA = \frac{f_{UE}}{f_{SO}}$, the ratio between the total system cost of the user equilbrium and the system optimal - it is a measure of the system penalty for selfish free will ($f_{UE} > f_{SO}$), i.e. if people choose what's best for them not what is best for the system. The reason for this is that the user equilibrium finds whats fair for everyone and the system optimal finds what is best for the system (often these don't match up).
- To find the User Equilibrium, we need to solve a **complementarity** problem (i.e. go through all the possible scenarios). However, beautifully, we can also find it using the equivalent **Beckmann** formulation: Minimise $\hat{f} := \int_0^{x_1} c_1(\bar{x})d\bar{x} + \int_0^{x_2} c_2(\bar{x})d\bar{x} + ...$ subject to $x_1 + x_2 + ... + x_n = d$ and $x_1, x_2, ..., x_n \geq 0$.
- Summary:
  - To find SO problem: Minimise $f = \sum_{i=1}^{n} x_i c_i(x_i)$ given constraints
  - To find UE problem: Minimise $\hat{f} = \sum_{i=1}^{n} \int_0^{x_i} c_i(\bar{x})d\bar{x}$ or compute the complementarity problem: For each subset $I := \{i_1, i_2, ..., i,\}$ or $\{1, 2, ..., n\}$, solve $c_{i1}(x_{i1}) = c_{i2}(x_{i2}) = ... = c_{im}(x_{im}) = C$ and check $c_i(0) \geq C$ for all $i \notin I$
- Now we get to more advanced networks, so we introduce y_i, which represents the **route flows**. The **link flows** will be a function of **route flows**. It like we distinguish between the possible route choices someone can take (y) and then consider who, given all the routes, will end up on a particular road (x). Note that, like for x, $y_1 + y_2 + ... + y_n = d$ and $y_i \geq 0$.
- Braess network - A counter intuitive network, whereby removing a short-cut actually improves traffic.
- You can create a **link-route incidence matrix** A, such that $\mathbf{x} = A\mathbf{y}$, where $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ (link flows, link variables), $\mathbf{y} = [y_1, y_2, ..., y_n]$ (route flows, route variables). However, one cannot (uniquely) find route flows from link flows - so UE and SO problems have a unique solution in link flows, but not usually for route flows. System cost f and Beckmann function $\hat{f}$ are functions of link flows x, so one could minimise either $f(A\mathbf{y})$ or $\hat{f}(A\mathbf{y})$.
- **Solving Assignment Problems using link variables**
  - It is possible to solve assignments without ever computing the routes or route flows.
  - Simply create equations of input/output flows for all junctions/nodes and put them into a matrix formulation $B\mathbf{x} = \mathbf{s}, \mathbf{x} \geq 0$ and then minimise $f(\mathbf{x})$ or $\hat{f}(\mathbf{x})$.
- The classical Wardrop and UE assumptions that all users have perfect information, are perfectly rational and all behave the same, is simply too unrealistic. A new modelling idea (**Stochastic User Equilibrium (SUE)**) adds uncertainty and randomness, thus enabling us to model imperfect information, irrationality and non-homogenous route choices.
  - Use **SoftMax** function with β parameter (as a sensitivity scalar), as a means assigning proportions $p_i = \frac{x_i}{d}$ of users making a particular choice. Note that $\sum_{i=1}^{n} p_i = 1$. Specifically the softmax function is: $p_i = \frac{f(c_i)}{\sum_{j=1}^{n} f(c_j)}$, where $f(c) = e^{-\beta c}$. Sometimes choice models are expressed in terms of **utility** rather than cost, $u_i = -c_i$.
  - Compute route costs c_i in terms of route flows y_i and then solve $\frac{y_i}{d_k} = f(c_i)/\sum_{j=1}^{n} f(c_j)$
  - Check out his matlab code, if you need to for coursework.

3 Microscopic Modelling

- explicitly model behaviour dynamics of every particle (vehicle, pedestrian, etc...)
- **Crowd/ Pedestrian Modelling**. Important for improving efficiency and safety in public spaces. It can be used to evaluate a new design before being built, or evaluate a proposed change or to analyse and improve understanding of an existing problem. Identify bottleneck areas (high density crowds).
  - Different modelling techniques: Cellular-automata, Social-force models, Queuing-network models, Fluid-

dynamic models (from more microscopic and discrete to more macroscopic and continuous).
  - Highly visually realistic models: PLEdestrians, Heuristic based crows modelling.
  - **Social Force models**:
    * Based off Newton's second law, idea is that pedestrians are driven by 'social forces'. These types of models are continues both in space and time.
    * Pedestrians are attracted to a particular goal (attraction force) and repelled by other pedestrians and obstacles (repelling force).
    * The model is:
$$\frac{d\mathbf{v}_\alpha(t)}{dt} = \mathbf{f}_\alpha(t) + \boldsymbol{\xi}_\alpha(t),$$
$$\mathbf{f}_\alpha(t) = \frac{1}{\tau_\alpha}(\mathbf{v}^0_\alpha \mathbf{e}_\alpha - \mathbf{v}_\alpha) + \sum_{\beta(\neq\alpha)} \mathbf{f}_{\alpha\beta}(t) + \sum_i \mathbf{f}_{\alpha i}(t).$$
      where $\frac{1}{\tau_\alpha}$ is the acceleration time,$\mathbf{v}^0_\alpha \mathbf{e}_\alpha$ is the desired velocity, $\mathbf{v}_\alpha$ is the actual velocity, $sum_{\beta(\neq\alpha)}\mathbf{f}_{\alpha\beta}(t)$ are the forces from other pedestrians, and $\sum_i \mathbf{f}_{\alpha i}(t)$. are the forces from obstacles. The pedestrian force $\mathbf{f}_{\alpha\beta}(t) = A_p \frac{\mathbf{d}}{||\mathbf{d}||} e^{-||\mathbf{d}||/B_p}$, A_p is the interaction strength, and B_p is the interaction range - you can also add a social force component $A_s \frac{\mathbf{d}}{||\mathbf{d}||} e^{-||\mathbf{d}||/B_s}$ to add a weaker longer range social force. The boundary force $\mathbf{f}_{\alpha i}(t) = A \frac{\mathbf{d}}{||\mathbf{d}||} e^{-||\mathbf{d}||/B}$, where $\mathbf{d}$ is the vector pointing from a boundary i to pedestrian α, again A, B are the interaction strength and range parameters.
    * An elliptical social-force model, is an improved version of the social force model?
    * To achieve better navigation, a potential field can be created. The negative gradients of the potential field give the direction leading the destination.

- Stop-and-go waves break into 'crowd turbulence'.
- The pedestrian 'fundamental diagram' is the relationship between crowd density and crows flow (average speed). Crowd flow equals crowd density times crowd speed.
- Use real data and compare these to simulations to optimize model parameters.
- **Self-organisation phenomena**. For bi-directional flows, *lane formation* of uniform walking direction will spontaneously emerge. Can use this to validate a model (amongst others, see last slide of microsim2 lecture).
- Adding obstacles can increase flow.
- **Exam based questions:**
  - **Autonomous cars**
    * See the Optimal Velocity Model (Bando et. al, 1995).
    * Capacity
  - **Social behaviours in crowds**
    * Social-force model (see above). You can add an angular and distance component.
    * Look at Anders's code.
  - **Evacuation model**
    * Desired direction of motion $\mathbf{e} = \frac{\mathbf{d}-\mathbf{x}}{||\mathbf{d}-\mathbf{x}||}$, where $\mathbf{d} = [d_1, d_2]$ the position of the exit, and $\mathbf{x}$ is the pedestrians position.
    * Can add multiple exits and then add conditions.
    * For more complex structures, we can generate potential fields.
  - **Modelling spread of disease in Crowds**
    * Just the SIR (susceptible, infected, removed) model.

4 Macroscopic Modelling

- Model population behaviour (zoomed out) - like a fluid. Valid when you have large number particles (vehicles) and flow features have longer length scale than particle (vehicle) separation. Macroscopic approach is okay for traffic flow on strategic highway (motorway). Motivation: Smart Motorways - Active Traffic Management.
- Sensor technologies: Inductance loops, count passing cars by sensing changes in inductance (with two, can compute speed, acceleration etc...); Average speed cameras. Bluetooth, GOS, Radar, CCTV, dashcams. Each sensor retrieves a different kind of data.
- MIDAS system - multiple loops on highways giving one minute reporting.
- Sometimes optimising flow and speed don't work together.
- You get 'wave' effects, known as phantom traffic jams.
- Key macroscopic variables: average speed u or v (Length/time), density ρ or k (vehicles/km), flow q (vehicles per time past a point), and occupancy $\mathcal{O}$ (proportion of road length covered by vehicles - dimensionless number).
- Equations: $q = \rho u$, $u = u_f(1 - \frac{\rho}{\rho_j}$, $q = u_f \rho(1 - \frac{\rho}{\rho_j}$ $q = \frac{2q_m u(u_f - u)}{u_f^2}$, where u_f is the free-flow speed, ρ_j is the jam density, q_m is the maximum flow, and if you plot these functions and find the maxima, then ρ_m, u_m are the density and speed at which maximum flow occurs. Jam density is where the velocity $u = 0$.
- Greenshields': 'Fundamental diagram' (FD) for speed flow: Diagram of speed vs flow. n

4.1 Space to time and vice versa

- Space-mean speed: Average speed $\bar{v}_S$ of vehicles/ particles viewed from above. Time-mean speed: Average speed $\bar{v}_T$ of vehicles/particles viewed from the road (e.g. from inductance loop), [He uses α_i for some of the v_i, why?, lecture 2, macroscopic]. Relationship between the two: From space-data to time-mean speed: $\bar{v}_S = \frac{1}{N}\sum_{i=1}^{N} v_i$, where v_i is the speed of vehicle i observed from above. $\bar{v}_T = \frac{\sum_{i=1}^{N} v_i^2}{\sum_{i=1}^{N} v_i}$ and their relationship: $\bar{v}_T = \bar{v}_S + \frac{\sigma_S^2}{\bar{v}_S}$, where $\sigma_S^2 = Var(v_1, v_2, ..., v_N)$. Vice versa, from time-data to space-mean speed: $\bar{v}_S = \frac{N}{\sum_{i=1}^{N}\frac{1}{v_i}}$, $\bar{v}_T = \frac{1}{N}\sum_{i=1}^{N} v_i$ and $\bar{v}_s \approx \bar{v}_T - \frac{\sigma_T^2}{\bar{v}_T}$, where $\sigma_T^2 = Var(v_1, v_2, ..., v_N)$ - unfortunately MIDAS data does not provide σ_T.
- Space to Time: $\bar{v}_S = 1/\int_0^\infty \frac{f(v)}{v} dv$ and $\bar{v}_T = \bar{v}_S + \frac{\sigma_S^2}{\bar{v}_S}$
- Time to Space: If we let $f(v)$ be the probability density funciton of speed of vehicles viewed by an inductance loop, then $\bar{v}_T = E(V) = \int_0^\infty v f(v) dv$, $\bar{v}_T = 1/\int_0^\infty \frac{f(v)}{v} dv$ and $\bar{v}_S \approx \bar{v}_T - \sigma_T^2/\bar{v}_T$ still holds, but now $\sigma_T^2 = \int_0^\infty v^2 f(v) dv - \bar{v}_T^2$
- Edie has some more generalised definitions in lecture 2 macroscopic, but only seems to be one slide so I'll skip them for now.

4.2 Continous Flow

- Density flow PDE (Lighthill - Whitham-Richards):

$$\frac{\partial \rho(x,t)}{\partial t} + \nabla \cdot (\rho(x,t)u(x,t)) = 0 \Rightarrow \frac{\partial \rho(x,t)}{\partial t} + \nabla \cdot (\rho(x,t)V(\rho)) = 0 \Rightarrow 1D \rightarrow = \frac{\partial \rho}{\partial t} + Q'(\rho)\frac{\partial \rho}{\partial x} = 0,$$

  where $Q(\rho) = \rho V(\rho)$.
- Solve using method of characterisitics.
- Goes into shocks, and cases where the solutions intersect and don't make sense.