NOTES TRAJETS TL, 2023

F. Bavaud

1 Notations

1.1 Lignes et flux

Les arrêts de bus $i = 1, \ldots, n$, formant l'ensemble V des noeuds du réseau, sont associés à l'une des lignes TL unidirectionnelles $l = 1, \ldots, m$.

On note par $l[i]$ la ligne sur laquelle est située l'arrêt i.

On note par $\alpha(l)$ l'arrêt de départ de la ligne l, et par $\omega(l)$ l'arrêt terminus de la ligne l.

On note par $F(i)$ (forward) l'arrêt suivant. La fonction est bien définie sauf $i = \omega(l[i])$ (i.e si i est le terminus de ligne), auquel cas $F[i] = \emptyset$.

On note par $B(i)$ (forward) l'arrêt précédent. La fonction est bien définie sauf $i = \alpha(l[i])$ (i.e si i est le départ de ligne), auquel cas $B[i] = \emptyset$.

Les flux de passagers x_{ij} de i vers j sont de deux types: (1) intra-ligne ssi $j = F(i)$ (ce qui implique $l[i] = l[j]$), noté alors y_{ij}; (2) inter-ligne (transfert péderstre) ssi $l[i] \neq l[j]$, noté alors z_{ij}. Comme un flux est soit de l'un, soit de l'autre type, on a

$$x_{ij} = y_{ij} + z_{ij}$$

2 Données

Les données sont constituées

(a) des montées a_i à l'arrêt i, avec $a_i = 0$ si $i = \omega(l[i])$ (i.e si i est un terminus de ligne)

(b) des descentes b_i à l'arrêt i, avec $b_i = 0$ si $i = \alpha(l[i])$ (i.e si i est un départ de ligne)

(c) du *tenseur d'intermédiarité* (betweenness tensor) χ_{ij}^{st}, défini comme

$$\begin{cases} \chi_{ij}^{st} = 1 & \text{si le plus court chemin de } s \text{ vers } t \text{ (supposé unique) passe par l'arrête } ij, \\ \chi_{ij}^{st} = 0 & \text{sinon.} \end{cases} \tag{1}$$

Il possède n^4 composantes. Dans un second temps, on pourra relaxer l'hypothèse d'un chemin le plus court unique de s à t, et remplacer χ_{ij}^{st} par une fonction (fortement) pénalisée par la longueur du chemin.

Le support S_Z des arrêtes ij de transfert, i.e. l'ensemble des arrêtes ij pour lesquelles il est possible que $z_{ij} > 0$, est tel que

$$\max_{st} \chi_{ij}^{st} = 1$$

i.e. si l'arrête ij est située sur un plus court chemin d'au moins une entrée s et une sortie t du réseau.

Les transferts possibles $z_{ij} > 0$ impliquent que i et j soient suffisamment proches, i.e. appartiennent à un même *super-stop* ou *crossing* $c = 1, \ldots, r$. On note par V_c ($c = 1, \ldots, r$) l'ensemble des noeuds appartenant au super-stop c. Un noeud i ne peut appartenir à deux super-stops, mais il peut n'appartenir à aucun super-stop. On note par V_0 l'ensemble des noeuds n'appartenant à aucun super-stop. Par construction, V est partitionnné comme

$V = \{V_0, V_1, \ldots, V_r\}$. Aussi, $z_{ij} > 0$ n'est possible que si i et j appartiennent au même super-stop (et $i \neq j$) et ainsi

$$S_Z = \bigcup_{c=1}^{r} (V_c \times V_c) \quad \setminus \quad \bigcup_{i=1}^{n} (\{i\} \times \{i\})$$

Le support S_Y des arrêtes ij intra-ligne, i.e. l'ensemble des arrêtes ij pour lesquelles il est possible que $y_{ij} > 0$, est constitué des paires telles que $j = F(i)$.

2.1 Proportion de transferts

On a aussi pour tout i qui n'est ni terminus ni départ.

$$y_{i,F(i)} = y_{B(i),i} + a_i - b_i \tag{2}$$

ce qui permet de déterminer les flux intra-ligne y_{ij}.

Parmi les passagers montant en i, certains viennent d'une autre ligne, d'autres entrent dans le réseau:

$$a_i = z_{\bullet i} + n_{i\bullet} \tag{3}$$

Parmi les passagers descendant en i, certains commutent vers une autre ligne, d'autres sortent du réseau:

$$b_i = z_{i\bullet} + n_{\bullet i} \tag{4}$$

Naturellement

$$a_\bullet = b_\bullet = z_{\bullet\bullet} + n_{\bullet\bullet}$$

La quantité $n_{\bullet\bullet}$ compte le nombre de personnes utilisant le réseau. La quantité $z_{\bullet\bullet}$ compte le nombre total de transferts (zéro, un, deux, trois... transferts par personne) effectués par les utilisateurs du réseau. La proportion moyenne de transferts par utilisateur est $z_{\bullet\bullet}/n_{\bullet\bullet}$.

2.2 Flux compatibles

Les flux $\mathbf{X} = (\mathbf{Y}, \mathbf{Z})$ sont compatibles avec les montées-descentes $(\mathbf{a}, \mathbf{b})$ ssi

$$y_{i,F(i)} = y_{B(i),i} + a_i - b_i \qquad\qquad z_{\bullet i} \leq a_i \qquad\qquad z_{i\bullet} \leq b_i \tag{5}$$

3 Inférence

3.1 Identité de GG et centralité

On considère toutes les paires possibles $(s, t) \in V^2$, où s est l'arrêt (le noeud) d'entrée d'un passager dans le réseau, et t est l'arrêt (le noeud) de sortie de ce passager dans le réseau. Il suffit de considérer $s \neq t$. On veut estimer n_{st}, le nombre de passagers entrant dans le réseau en s, et sortant du réseau en t. On a l'identité de GG

$$x_{ij} = y_{ij} + z_{ij} = \sum_{st} \chi_{ij}^{st}\, n_{st} \tag{6}$$

C'est seulement par la quantité χ_{ij}^{st} que les trajectoires z_{ii} ou comme z_{ii-} sont interdites (section 3.1.1).

On a $x_{ij} = y_{ij} + z_{ij}$, où les deux termes ne peuvent être non-nuls simultanément: y_{ij} est possiblement positif ssi $j = F(i)$, et z_{ij} est possiblement positif ssi $i, j \in I$ (même jonction) avec $i \neq j$.

Un chemin (d'utilisation du réseau) st commence toujours par une arrête intra $(s, F(s))$, et se termine toujours par une arrête intra $(B(t), t)$. Des arrêtes de transferts peuvent être empruntées dans les parties *intermédiaires* du chemin. On définit

$$\begin{cases} \gamma_i^{st} := 1 & \text{si le plus court chemin de } s \text{ vers } t \text{ passe par le stop } i \ (i = s \text{ et } i = t \text{ possibles}), \\ \gamma_i^{st} = 0 & \text{sinon.} \end{cases} \tag{7}$$

Alors

$$\gamma_i^{st} = \chi_{i\bullet}^{st} + \delta_{it} = \chi_{\bullet i}^{st} + \delta_{is}$$

On pose $\Gamma_i := \sum_{st} \gamma_i^{st}\, n_{st}$, qui est un indice de *centralité d'intermédiarité* du stop i. En sommant (6), on a

$$y_{i\bullet} + z_{i\bullet} = \sum_{st} \gamma_i^{st}\, n_{st} - \sum_{s} n_{si} = \Gamma_i - n_{\bullet i} \tag{8}$$

De même,

$$y_{\bullet i} + z_{\bullet i} = \sum_{st} \gamma_i^{st}\, n_{st} - \sum_t n_{it} = \Gamma_i - n_{i\bullet} \tag{9}$$

Mais $b_i = z_{i\bullet} + n_{\bullet i}$, et donc (8) s'écrit

$$\Gamma_i = y_{i\bullet} + b_i = y_{i,F(i)} + b_i \tag{10}$$

De même, $a_i = z_{\bullet i} + n_{i\bullet}$, et donc (9) s'écrit

$$\Gamma_i = y_{\bullet i} + a_i = y_{B(i),i} + a_i \tag{11}$$

En particulier, $y_{i,F(i)} = y_{B(i),i} + a_i - b_i$, comme il se doit.

Si i est le début de ligne, alors $y_{B(i),i} = 0$ et $\Gamma_i = a_i$. Si i est le terminus, alors $y_{i,F(i)} = 0$ et $\Gamma_i = b_i$.

3.1.1 (*) Cas possiblement déviants

Pour z_{ii} : on a que $\chi_{ii}^{st} = 0$: l'"arrête" ii serait empruntée par un voyageur descendant en i, puis remontant en i (par le/un bus d'après), possible mais certainement pas le plus court chemin: irrationnel. Sauf si on hésite à poursuivre et qu'on décide quand même de poursuivre après réflexion. Ou que, voyageur sans billet, des contrôleurs s'apprêtent à monter en i, on y descend précipitamment et attend le prochain bus: finalement très rationnel, possiblement...

Pour z_{ii-} : passer à l'arrêt opposé (on suppose que $B(i)- = F(i-)$ (i.e. que les lignes étaient parallèles "avant l'arrêt i") ne peut pas être le plus court chemin: si le voyageur descend à i, il a passé au moins par $B(i)$, et s'il monte en $i-$, il passe au moins par $F(i-)$, Or, quel que soient s et t,

$$d_{sB(i)} + d_{B(i)i} + d_{ii-} + d_{i-F(i-)} + d_{F(i-)t} \;>\; d_{sB(i)} + d_{B(i)B(i)-} + d_{F(i-)t}$$

car

$$d_{B(i)i} + d_{ii-} + d_{i-F(i-)} \;>\; d_{B(i)B(i)-}$$

où d_{ii-} est le coût de traverser la route en i, et $d_{B(i)B(i)-}$ celui de traverser la route l'arrêt d'avant. Mais s'il est épouvantable de traverser la route en i (camions, pas de passage sécurisé), ou s'il est embarrassant de traverser la route en i (par exemple sous les yeux d'une connaissance possiblement dans les parages qui nous croyait à l'étranger, ou en arrêt maladie), alors de nouveau il peut être préférable de descendre un arrêt plus loin puis de rebrousser chemin: de nouveau, finalement très rationnel, possiblement...

3.2 Independence assumption on stops

Assume $n_{st}^{(r)} = \frac{n_{s\bullet}^{(r)} n_{\bullet t}^{(r)}}{n_{\bullet\bullet}^{(r)}}$. Compute $\mathbf{X}^{(r+1)} = \boldsymbol{\chi}\mathbf{N}^{(r)}$, then compute

$$n_{i\bullet}^{(r+1)} = a_i - z_{\bullet i}^{(r+1)} \qquad n_{\bullet i}^{(r+1)} = b_i - z_{i\bullet}^{(r+1)} \qquad n_{st}^{(r+1)} = \frac{n_{s\bullet}^{(r+1)} n_{\bullet t}^{(r+1)}}{n_{\bullet\bullet}^{(r+1)}} \tag{12}$$

Start from some $n_{s\bullet}^{(0)}$ and $n_{\bullet t}^{(0)}$ with coinciding totals (such as $n_{s\bullet}^{(0)} = a_s$ and $n_{\bullet t}^{(0)} = b_t$, i.e. no transfers $\mathbf{Z}^{(0)} = \mathbf{0}$), and iterate (12) until convergence.

Il faudrait en plus s'assurer de la compatibilité $a_i \geq z_{\bullet i}^{(r+1)}$ et $b_i \geq z_{i\bullet}^{(r+1)}$ à chaque étape.

3.2.1 *** contre-argument? Mais non***

*** Mais en fait $a_i - z_{\bullet i}^{(r)} = n_{i\bullet}^{(r)}$, et donc la première identité (12) s'écrit $n_{i\bullet}^{(r+1)} = n_{i\bullet}^{(r)}$. De même, $b_i - z_{i\bullet}^{(r)} = n_{\bullet i}^{(r)}$ et donc $n_{\bullet i}^{(r+1)} = n_{\bullet i}^{(r)}$. En conséquence, toute solution initiale est déjà un point fixe, il n'y aucune évolution / convergence. Si on part avec une condition initiale sans transferts, toutes les lignes sont de facto déconnectées, avec des flux $n_{st}^{(0)}$ *purement mono-lignes*, et ainsi aucun transfert $z_{ij} > 0$ ne va jamais émerger de $\mathbf{X}^{(r)} = \boldsymbol{\chi}\mathbf{N}^{(r)}$ ... $\frown$. Il faudrait essayer travailler au niveau de *jonctions* aux flux indépendants.

3.3 Independence assumption on junctions

Each stop i belongs to a single line $\ell(i)$, with a successor stop $F(i)$ always defined unless i is the terminus of the line.

Also, each stop i belongs to a single super-stop junction $I[i]$, including “bare junctions" where no transfers take place. Define the edge-(junction-)path incidence matrix

$$\begin{cases} \chi_{ij}^{ST} = 1 & \text{si le plus court chemin (supposé unique) de la jonction } S \text{ vers la jonction } T \text{ passe par l'arrête } ij, \\ \chi_{ij}^{ST} = 0 & \text{sinon.} \end{cases} \tag{13}$$

Un chemin (d'utilisation du réseau) ST commence toujours par l'arrête intra $(i, F(i))$ avec $i \in S$, et se termine toujours par une arrête intra $(B(j), j)$ avec $j \in T$. Des arrêtes de transferts peuvent être empruntées dans les parties *intermédiaires* du chemin.

On indique, systématiquement, des jonctions par des majuscules, dont les indices *inférieurs* (i.e. pas comme dans χ_{ij}^{ST}) dénotent des agrégations (sommations) de stops dans les jonctions. Par exemple:

$$n_{ST} = \sum_{s \in S; t \in T} n_{st} \qquad y_{\bullet S} = \sum_{s \in S} y_{\bullet s} \qquad \Gamma_S = \sum_{i \in S} \Gamma_i$$

L'identité de GG devient

$$x_{ij} = y_{ij} + z_{ij} = \sum_{ST} \chi_{ij}^{ST}\, n_{ST} \tag{14}$$

On a $\chi_{ij}^{ST} \neq \sum_{s \in S; t \in T} \chi_{ij}^{st}$. L'identité (14) diffère de (6) qui s'écrit

$$x_{ij} = y_{ij} + z_{ij} = \sum_{st} \chi_{ij}^{st}\, n_{st} = \sum_{ST} \sum_{s \in S; t \in T} \chi_{ij}^{st}\, n_{st}$$

L'itération commence par le calcul de $\mathbf{X}^{(r+1)} = \boldsymbol{\chi}\mathbf{N}^{(r)}$ selon (14), puis par le calcul de l'analogue des identités (12) qui deviennent

$$n_{S\bullet}^{(r+1)} = a_S - z_{\bullet S}^{(r+1)} \qquad n_{\bullet T}^{(r+1)} = b_T - z_{T\bullet}^{(r+1)} \qquad n_{ST}^{(r+1)} = \frac{n_{S\bullet}^{(r+1)} n_{\bullet T}^{(r+1)}}{n_{\bullet\bullet}^{(r+1)}} \tag{15}$$

A noter que $z_{\bullet S} = z_{S\bullet}$ quel que soit $\mathbf{N}$. Ainsi, on a toujours $n_{S\bullet} - n_{\bullet S} = a_S - b_S$ (les flux $\mathbf{N}$ sont généralement marginalement inhomogènes), avec $a_\bullet = b_\bullet$ évidemment.

3.4 Une seule ligne

On veut trouver $\mathbf{N} = (n_{st})$ (qu'on peut normaliser et considérer comme une distribution de probabilité) qui maximise l'entropie sous les contraintes $n_{s\bullet} = a_s$ et $n_{\bullet t} = b_t$, i.e. $\sum_{st} n_{st}\delta_{si} = a_i$ et $\sum_{st} n_{st}\delta_{tj} = b_j$, avec le prior théorique $f_{st}^M = \frac{2}{l(l-1)} I(s < t)$. Avec les multiplicateurs de Lagrange (intégrant la normalisation / le total), la solution est de la forme

$$n_{st} = I(s<t) \prod_i \exp(-\lambda_i \delta_{si}) \prod_j \exp(-\mu_j \delta_{tj}) = I(s<t)\, c_s\, d_t \qquad \sum_{s<t} c_s\, d_t = n_{\bullet\bullet} = a_\bullet = b_\bullet$$

Les contraintes sont que

$$a_i = \sum_{st} I(s<t)\, c_s\, d_t\, \delta_{si} = c_i \sum_{t>i} d_t = c_i\, D_i \qquad D_i := \sum_{t>i} d_t$$

$$b_j = \sum_{st} I(s<t)\, c_s\, d_t\, \delta_{tj} = d_j \sum_{s<j} c_s = d_j\, C_j \qquad C_j := \sum_{s<j} c_s$$

Soit m_{st} le nombre de voyageurs partant de s et toujours présents en t, et soit ρ_j la probabilité de descendre en j. On a

$$n_{st} = m_{st}\rho_t \qquad n_{s,t+1} = m_{st}(1-\rho_t)\rho_{t+1} \qquad \frac{n_{s,t+1}}{n_{st}} = \frac{(1-\rho_t)\rho_{t+1}}{\rho_t} = \frac{d_{t+1}}{d_t}$$

Ainsi

$$\frac{d_{t+2}}{d_{t+1}} = \frac{(1-\rho_{t+1})\rho_{t+2}}{\rho_{t+1}} \qquad \frac{d_{t+2}}{d_t} = \frac{d_{t+2}}{d_{t+1}}\frac{d_{t+1}}{d_t} = \frac{(1-\rho_{t+1})\rho_{t+2}}{\rho_{t+1}}\frac{(1-\rho_t)\rho_{t+1}}{\rho_t} = \rho_{t+2}(1-\rho_{t+1})(1-\rho_t)\frac{1}{\rho_t}$$

$$\frac{d_{t+q}}{d_t} = \rho_{t+q}(1-\rho_{t+q-1})(1-\rho_{t+q-2})\dots(1-\rho_{t+1})(1-\rho_t)\frac{1}{\rho_t}$$

$$d_l + d_{l-1} = d_2[\rho_l(1-\rho_{l-1}) + \rho_{l-1}](1-\rho_{l-2})(1-\rho_{l-3})\dots(1-\rho_3)(1-\rho_2)\frac{1}{\rho_2}$$

$$D_i = \sum_{t>i} d_t$$

Soit z_{st} la proportion de voyageurs présents dans le bus en s et toujours présents en $t > s$. Soit τ_s la probabilité de monter en s. Alors $f_{st} = \tau_s z_{st} \rho_t$

4 *** OLD JUNK

4.1 Une seule ligne

On ordonne les arrêts $i = 1, \ldots, M$ de façon croissante dans le sens du parcours (i.e. $F(i) = i + 1$). Soit $A_i = \sum_{j \leq i} a_j$, respectivement $B_i = \sum_{j \leq i} b_j$, le nombre cumulé de montées, respectivement de descentes, jusqu'en i. Il faut que $A_i \geq B_i$ et $a_\bullet = b_\bullet$.

Alors $y_{i,F(i)} = A_i - B_i$, et $y_{B(i),i} = (A_i - a_i) - (B_i - b_i) = y_{i,F(i)} - a_i + b_i$, qui satisfait (2) comme il se doit.

On a $\chi_{ij}^{st} = I(j = i+1)\, I(s < t)\, I(s \leq i)\, I(j \leq t) = I(j = i+1)\, I(s \leq i)\, I(j \leq t)$

Soit $n_{st} = c_s d_t I(s < t)$. En substituant dans (6), on a

$$y_{ij} = I(j = i+1) \sum_{s \leq i} \sum_{t \geq j} c_s d_t$$

4.1.1 Maxent

Let $f_{st}^D := \frac{n_{st}}{n_{\bullet\bullet}}$ be the empirical distribution, incompletely observed: only $n_{s\bullet} = a_s$ and $n_\bullet = b_t$ are known (entailing $n_{\bullet\bullet} = a_\bullet = b_\bullet$). Consider the "semi-uniform" prior $f_{st}^M = \frac{I(s<t)}{\frac{1}{2}M(M-1)}$ (where M is the number of stops of the single line), assigning a constant probability to any st-trip provided $s < t$.

The MaxEnt solution $\tilde{f}_{st}^D$ minimizing $K(f^D || f^M) = \sum_{st} f_{st}^D \ln \frac{f_{st}^D}{f_{st}^M}$ under the constraints $\sum_t f_{s't}^D = \sum_{st} f_{st}^D I(s = s') = \frac{a_{s'}}{a_\bullet}$ and $\sum_s f_{st}^D = \sum_{st'} f_{st'}^D I(t = t') = \frac{b_t}{b_\bullet}$ is of the form

$$\tilde{f}_{st}^D = \frac{f_{st}^M \exp(\phi_s) \exp(\gamma_t)}{Z}$$

that is

$$\tilde{n}_{st}^D = n_{\bullet\bullet} \tilde{f}_{st}^D = c_s d_t I(s < t) \qquad \text{avec} \quad \sum_{st} c_s d_t I(s < t) = n_{\bullet\bullet}$$

with the constraints

$$c_s \sum_{t > s} d_t = a_s \qquad d_t \sum_{s < t} c_s = b_t \tag{16}$$

On a $a_M = 0$ et $b_1 = 0$. Aussi, pour tout $t = 1, \ldots, M$,

$$A_t := \sum_{s=1}^{t} a_s \quad \geq \quad B_t := \sum_{s=1}^{t} b_s$$

4.1.2 Interprétation markovienne de GG

On peut monter que la solution MaxEnt est de la forme

$$\tilde{n}_{st} = I(s < t)\, a_\bullet\, p_s^{\text{in}} c_{s+1} \ldots c_{t-1} p_t^{\text{out}} \qquad \tilde{n}_{s,s+1} = a_\bullet p_s^{\text{in}} p_{s+1}^{\text{out}} \tag{17}$$

avec $p_s^{\text{in}} = \frac{a_s}{a_\bullet}$, $\quad p_t^{\text{out}} = \frac{b_t}{y_{t-1,t}}$ et $c_i = 1 - p_i^{\text{out}}$. On a $p_M^{\text{out}} = \frac{b_M}{y_{M-1,M}} = 1$ (terminus, tout le monde descend).

*** preuve ***

4.1.3 Identité de GG

4.1.4 Maxent for multilines

A *junction* I (not "superstop") is a set of disjoint set I of stops "not too far apart", on which transfers can take place. Any transfer ij is allowed for $i, j \in I$, with the exception of $i = j$ (unboarding at i followed by boarding at the same stop) and $j = i_-$, where i_- denotes the stop "opposite to i", that is crossing the road to take the line opposite to

$n_{\bullet\bullet}$ is not known. Its maximal value is reached under no transfers, that is when each boarding is an entrance in the network, and when each unboarding an exit. That is

$$n_{\bullet\bullet} \leq a_\bullet = b_\bullet$$

On the other direction, $n_{\bullet\bullet}$ is minimum when transfers are maximal: that is, one would like, for every super-stop I to match each unboarding traveller with a boarding traveller, leaving I traveller-free. Yet superstop total boardings and unboardings do not match in general:

$$a_I = \sum_{i \in I} a_i \qquad b_I = \sum_{i \in I} b_i \qquad a_I \neq b_I \text{ in general.}$$

By construction

$$b_I = z_{I\bullet} + n_{\bullet I} \qquad a_I = z_{\bullet I} + n_{I\bullet} \tag{18}$$

But $\mathbf{Z}$ is a directed sum of flows concentrated on junctions I, and zero outside junctions (denoted 0), i.e. $\mathbf{Z} = \mathbf{Z}^0 \bigoplus_I \mathbf{Z}^I$. Hence

$$z_{I\bullet} = z_{\bullet I} \qquad n_{I\bullet} - n_{\bullet I} = a_I - b_I$$

Note that

$$a_\bullet = a_0 + \sum_I a_I = b_\bullet = b_0 + \sum_I b_I \tag{19}$$

where a_0 is the number of travellers boarding outside junctions, and b_0 is the number of travellers un-boarding outside junctions.

A maximum amount of $\min(a_I, b_I)$ can thus be transferred at (supstop) junction I, and

$$n_{\bullet\bullet} \geq a_\bullet - \sum_I \min(a_I, b_I) = \frac{1}{2}(a_0 + b_0 + \sum_I |a_I - b_I|) \tag{20}$$

Constraints such that no transfers between opposite directions take place are not taken into account, so the lower bound (20) is too small in some cases (when the activity on the other lines is too small).

Geography defines junctions as disjoint set of stops "not too far apart", and forbids transfers between opposite directions by imposing that transfer $i \to j$ must necessarily take place on shortest st paths, that is edge ij belongs to the st shortest path (symmetric in composition for bidirectional lines), that is iff $\chi_{ij}^{st} = 1$. But there is no z_{st} such that $\chi_{ij}^{st} = 1$: if i and j are "opposite stops": boarding on s "before i" (some reference direction) and un-boarding at t "after j" is never on the shortest path, unless $s = i$ and $j = t$, in which case there is no transfer.

*** pas de transferts possibles: fantaisie de lignes de bus entièrement disconnectées...

Prior: $f_{st}^M = \frac{1}{z} I(t \neq s) I(t \neq s_-)$, where s_- is the bus stop "opposite" to s: one assumes twin, bidirectional bus lines). In particuler, one assumes that each stop can be reached

As a consequence, each bus top s (belonging to a junction of not) possesses an unique "opposite" bus stop s_-. Hence $z = \sum_{st} I(t \neq s) I(t \neq s_-) = M(M-2)$, where M is the total number of bus stops, namely

$$M = \sum_{l=1}^{L} M_l$$

where M_l is the number of bus stops at line l, and the number of lines L is even (twin, bidirectional bus lines).

4.2 MaxEnt multilines, revisited

L'identité de GG s'écrit

$$x_{ij} = \sum_{st} \chi_{ij}^{st} n_{st} = \sum_{[st] \ni (ij)} n_{st} \tag{21}$$

où (ij) est une arrête (orientée) du réseau, $[st]$ le plus court chemin (orienté, supposé unique) allant de s à t. On note $(ij) \in [st]$ ou $[st] \ni (ij)$ le fait que le chemin $[st]$ passe par l'arrête (ij), i.e. que l'arrête (ij) est sur le chemin $[st]$. L'expression $\sum_{[st]\ni(ij)} \ldots$ est la somme sur tous les chemins $[st]$ passant par (ij) (fixé).

La solution MaxEnt pour n_{st} (convertible en une probabilité jointe) avec un prior uniforme et les contraintes (21) (convertibles en moyennes) est

$$n_{st} = \prod_{(ij)} \exp(-\gamma_{ij}\, \chi_{ij}^{st}) = \prod_{(ij)\in[st]} \phi_{ij} \tag{22}$$

i.e. un produit de termes consécutifs sur les arrêtes de $[st]$.

En particulier, si $(s, F(s)) \in [s,t]$ (i.e. $F(s)$ est sur la ligne amenant vers t depuis s), alors $\boxed{n_{st} = \phi_{(s,F(s))}\, n_{F(s),t}}$: vision intra-ligne, on reste dans le bus, on était dans le bus en s, on y était soit monté, soit resté depuis $B(s)$. En fait, chaque stop s appartient à une ligne, et donc $F(s)$ est toujours défini. Aussi, si quelqu'en monte en s, il reste au moins jusqu'en $F(s)$. S'il descend en s, il ne peut pas être entré dans le réseau en s.

Si en plus on assume l'indépendance $n_{st} = n_{\bullet\bullet}\rho_s\rho_t$, alors de ce qui précède $\rho_s = \phi_{(s,F(s))}\, \rho_{F(s)}$, et donc $\phi_{(s,F(s))} = \frac{\rho_s}{\rho_{F(s)}}$. Si l'on généralise à (toutes, de deux sortes?) les arrêtes, on a $\boxed{\phi_{ij} = \frac{c_i}{c_j}}$ et

By the way: on est en fait vraiment intéressé par n_{ST} (S, T = jonctions "vraies" ou stops isolés) plutôt que par n_{st}.

Let M be the number of edges (within + transfer), and let N be the number of stops (nodes); so N^2 is the number of paths. The $M \times N^2$ matrix $\boldsymbol{\chi} = (\chi_{(ij)[st]}) \equiv (\chi_{ij}^{st})$ is the *edge-path incidence matrix*. Identity 21) reads $\mathbf{X} = \boldsymbol{\chi}\mathbf{N}$. Identity (22) reads $\log n_{st} = \sum_{(ij)\in[st]} \log \phi_{ij} = -\sum_{(ij)} \gamma_{ij}\, \chi_{(ij)[st]}$, that is $\log \mathbf{N} = -\boldsymbol{\chi}^\top \log \boldsymbol{\Gamma}$

5 EM approach

Initial: no transfers? Minimize $n_{\bullet\bullet} K(f^D || f^M)$?

If n_{st} is given, then y_{ij} (observed) and z_{ij} (unobserved) follow from $x_{ij} = \sum_{st} \chi_{ij}^{st}\, n_{st}$

But $\mathbf{Y}$ must obey $y_{i,F(i)} = y_{B(i),i} + a_i - b_i$

If y_{ij} and z_{ij} are given, how to infer n_{st}?