Distribuição das viagens por tempo de percurso

Pesquisa O/D e a matriz origem/destino

Csaba Deák

Encomendado para o Nº 1 da Revista da EBTU, que foi extinta antes da revista ver a luz.

Agosto, 1988

Resumo/ introduçao

As amostras das pequisas de tráfego sâo insuficientes para se poder expandir os fluxos interzonais e assim se obter a matriz origem/destino. No entanto, as mesmas sâo amplamente suficientes para se levantar a curva de distribuiçâo das viagens segundo o tempo de percurso, e ainda, desagregados segundo classes de populaçâo, modo de transporte e motivo da viagem. Dessa maneira, uma funçâo impedância da distância pode ser calibrada com grande precisâo e posteriormente utilizada para produzir estimativas confiáveis dos fluxos interzonais.

1 Problemas da expansao direta da matriz o/d na amostra

As consideraçôes abaixo, sobre o método de obtençâo de uma estimativa dos fluxos interzonais --matriz origem-destino-- resultam dos trabalhos preparatórios da aplicaçâo de nova pesquisa origem/destino aglomeraçâo urbana de Sâo Paulo, em 1987.

Por ocasiâo da pesquisa anterior (1977) o zoneamento de tráfego era composto de 243 zonas e haviam sido aplicados da ordem de 120.000 questionários em 25.000 domicílios, resultando em 180.000 viagens levantadas (sobre um total de 21 milhôes, das quais 16 milhôes motorizados). A matriz O/D continha portanto 243x243= 59.049 elementos, para a estimativa dos quais dispunha-se da ordem de 150.000 (viagens motorizadas) dados amostrados, ou seja, em torno de 2,5 observaçôes por par origem/destino. Nessas condiçôes, a expansâo pura e simples da matriz amostrada mediante seus respectivos fatores de expansâo resulta em uma matriz de viagens "estimadas" que contém a esmagadora maioria de seus elementos, nulos, e o restante, números quasi-aleatórios. Para nâo mencionar a deagregaçâo adicional das viagens segundo classes de populaçâo, modo de transporte, motivo e hora de realizaçâo e finalmente, motivos de viagem.

O quadro abaixo exemplifica a situaçâo, reproduzindo os elementos de uma linha da matriz, no caso, as viagens estimadas a partir de uma zona de tamanho médio (ZT 141:Vila Matilde; populaçâo: 102.000), situada na coroa intermediária da estrutura urbana. As viagens referem-se à faixa de renda de

5 a 12 salário mínimo (que agrega perto de 30% da populaçâo de Sâo Paulo e totaliza 43.000 em Vila Matilde), com destino no emprego e realizadas na hora-pico. A referida linha contém 8 elementos nâo-nulos (e portanto, 234 nulos, nâo assinalados). Para completar a informaçâo, estâo indicados os possíveis motivos a favor e contra a realizaçâo de viagens a cada destino, a saber, o número de empregos e a distância à origem. Os números falam por si, exibindo um caráter errático do número de viagens estimadas em relaçâo quer seja à distância, quer aos empregos, das zonas de destino. Mencione-se

Viagens externas com origem em V.Matilde
Faixa de renda 3, destino no emprego, hora pico

.	Destino	viagem	dist	Emprego
Nº	Nome	(Nº)	(km)	.
32	Belenzinho	848	6	21.167
17	Aclimaçâo	396	12	10.473
24	Consolaçâo	171	13	66.871
117	Ibirapuera	97	19	31.552
3	Brás	87	10	20.659
107	Santo André	20	12	10.033
102	Penha	4	2	18.777
171	Cumbica	3	13	23.891
.	Total	1.626	.	.

adicionalmente apenas que nâo foi estimada nenhuma viagem a qualquer das zonas de 1 a 16 que compôem o centro da cidade e totalizam 550.000 empregos, ou perto de um quinto do total de empregos da Regiâo Metropolitana.

2 A distribuiçao das viagens segundo classes de distância

Uma abordagem alternativa à obtençâo do padrâo de deslocamentos na área urbana e em última instância, da própria matriz O/D, passa por uma análise dos dados colhidos em pesquisa de tráfego nâo puramente estatística, senâo informada por uma idéia de modelo do comportamento dos deslocamentos. Note-se que na realidade, proceder à estimativa da matriz O/D *sem* se valer de um modelo de atrito da distância é algo similar a, por exemplo, estudar a trajetória de um projétil sem levar em conta as leis de Newton, vale dizer, ignorando as características mais elementares do processo em questâo. No caso dos deslocamentos, seria ignorar o fato elementar que viagens mais curtas sâo geralmente preferidas a viagens mais longas (e equidistantes sâo indiferentes), noçâo essa à base dos modelos gravitacionais, como também do modelo de enfoque probabilístico sumariamente descrito no que segue.

Segundo esse modelo, a probabilidade de uma viagem se realizar é uma funçâo da distância que separa as oportunidades de origem e de destino. Assim, a probabilidade de viagens casa-trabalho --para tomar um exemplo, que será utilizado daqui em diante-- de uma origem **i** com populaçâo **P**i, para uma zona de destino **j** com **E**j empregos a uma distância t_{ij} é proporcional a

$$v_{ij} = f(t_{ij})\cdot P_i\cdot E_j \ , \quad (1)$$

onde **f(t)** é uma funçâo da distância (essa última, designada por ***t*** porque é melhor expressa pelo tempo de percurso). Sendo entâo

$$f(t_{ij}) = \frac{v_{ij}}{P_i\cdot E_j} \ , \quad (2)$$

obtém-se a funçâo distância $f(t)$ levantando-se simplesmente a curva de frequência das oportunidades de viagem realizadas --isto é, cada viagem dividido por $P_i\cdot E_j$-- segundo o tempo de percurso, *independentemente* de sua origem e destino. Como já mencionado, a frequência das viagens amostradas fornece uma estimativa extremamente confiável das frequências efetivas, devido ao tamanho generoso da amostra, permitindo também, e ainda com folga, sua desagregaçâo segundo modo de transporte (individual, coletivo etc.) e faixas de renda (da populaçâo), como no exemplo que segue adiante. Na prática, a funçâo distância será "discretizada", isto é, obtida na forma de um histograma de passo suficientemente pequena para se

obter uma precisâo razoável, da ordem de 3 a 10 minutos. Assim, se o passo for D**t, as ordenadas discretas da curva de frequência serâo**

$$f(t_k) = \sum_i \sum_j \frac{u^k_{ij}}{P_i\cdot E_j} \ , \ k = 1,2,... \quad (3)$$

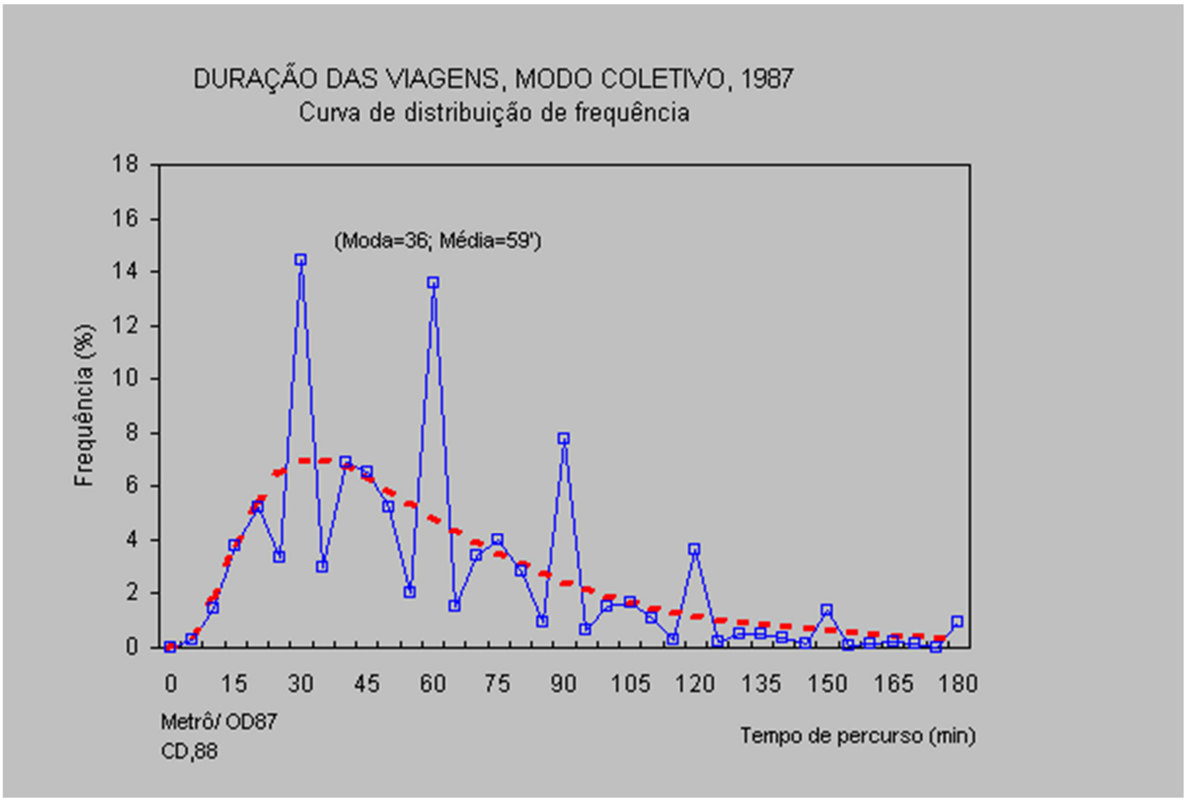
onde

$t_k = k\cdot\Delta t - \Delta t/2$, isto é, o ponto médio do k-ésimo intervalo de tempo,

$u^k_{ij} = 1$ se $t_k - \Delta t/2 < t_{ij} < t_k + \Delta t/2$, isto é, para as viagens cujo tempo de percurso está compreendido no k-ésimo intervalo, e

$u^k_{ij} = 0$ no caso contrário.

A *Fig.1* ao lado representa a distribuiçâo da frequência das viagens realizadas por transporte coletivo segundo o tempo de percurso em intervalos de 5 em 5 minutos até 2½ horas (150 minutos, valor para o qual a frequência acumulada atinge 98%), como levantadas na amostra da pesquisa O/D de 1977 em Sâo Paulo.



3 Calibraçao do modelo

O caráter disparatado das frequências 'observadas' deve-se ao arredondamento dos tempos de percurso declarados. A correçâo meramente visual e "a mâo livre" das mesmas resulta na curva tracejada, que já permite uma boa visualizaçào da distribuiçâo, assim como a escolha do tipo de funçâo algébrica a ser calibrada, no caso, uma distribuiçâo log-normal, da forma

$$\mathbf{f}(\mathbf{t}) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp \frac{-[\ln(\mathbf{t})-\overline{\ln(\mathbf{t})}]^2}{2\boldsymbol{\sigma}^2} \quad \mathbf{(4)}$$

ou seja, uma distribuiçâo normal onde em abscissa fica o logaritmo do tempo de percurso. Os dois parâmetros ln(**t**) e $\boldsymbol{\sigma}$ (média e desvio padrâo, respectivamente, do *logaritmo* dos tempos de percurso) dessa funçâo poderiam ser obtidos analiticamente ao preço de se levantar a distribuiçâo de frequência das viagens sobre uma base em que os tempos de percurso tenham sido previamente convertidos em seu próprio logaritmo. A alternativa aqui descrita é possivelmente menos elegante, porém permite um maior controle no processo de ajuste, particularmente das grandes proporçôes entre viagens curtas e longas. A calibraçâo é efetuada mediante uma regressâo linear combinada a um ajuste "manual", segundo os passos que seguem (cf. também *Fig.2* abaixo).

[1] Regressâo linear: Tomando-se a média modal, isto é, o tempo de percurso mais frequente, da própria amostra, como **t, efetua-se a regressâo sobre a funçâo (1) linearizada em**

$$\ln[f(t)] = \ln(c_o) + c_1 \cdot [\ln(t) - \ln(t)]^2 \quad (5)$$

utilizando-se os valores das frequências observadas $f_o(t)$, para se obter os parâmetros c_o e $c1$.

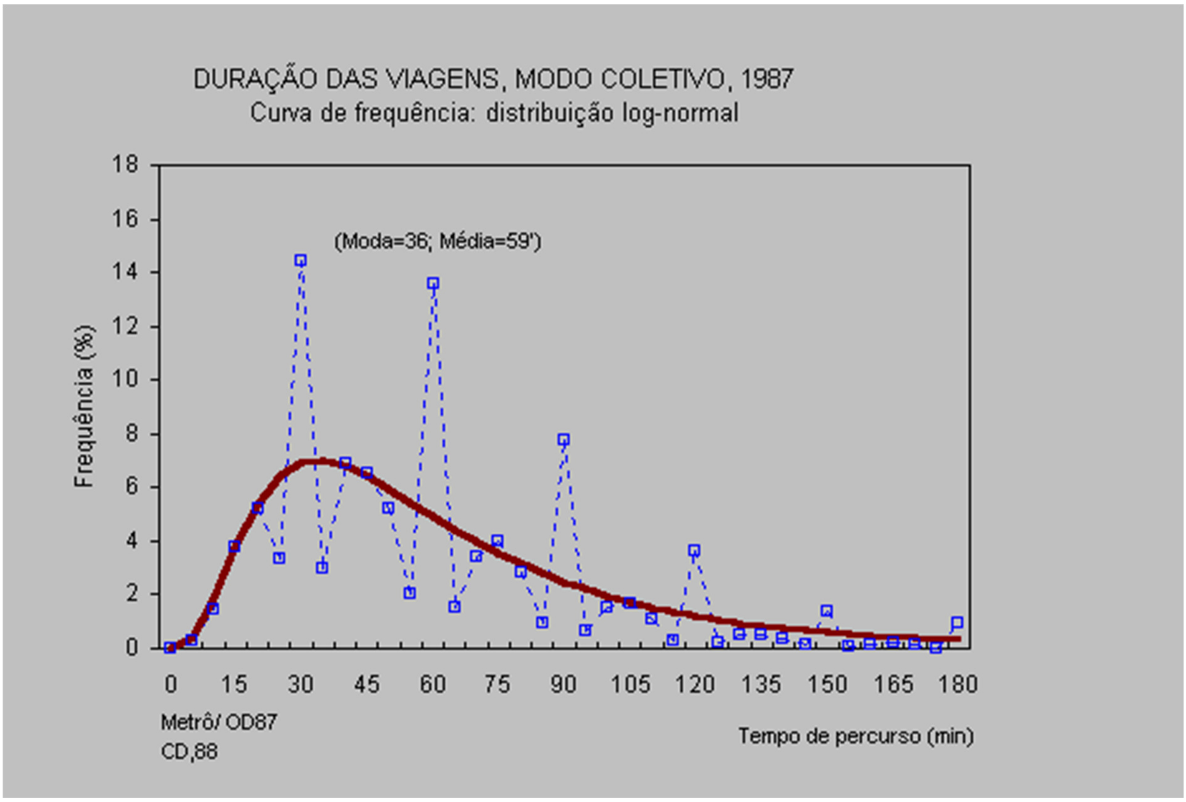
[2] Primeira estimativa: Obtém-se uma primeira estimativa da distribuiçâo de frequências $f'(t)$ utilizando-se (5).

[3] Normalizaçâo: Normaliza-se a distribuiçâo obtida de maneira que as frequências somem 1 (ou 100%), fazendo

$$f(t_k) = f'(t_k) \cdot 1/ \sum^{k} f'(t_k)$$

[4] Calibraçâo das viagens médias: Calculam-se as proporçôes de viagens médias q_o e q observadas e estimadas, respectivamente, da forma

$$q = \sum_{k=m}^{n} f(t_k)$$



com **m** e **n** escolhidos de maneira a compreender as duraçôes de viagem de 75 a 120 minutos, por exemplo. A seguir procura-se o valor de **t** que minimize a diferença entre **q**o e **q**, através de tentativas sucessivas, dando-se acréscimos sucessivos ao valor inicial de **t** no passo **[1]** e repetindo para cada novo valor os passos **[1]-[4]**. Os acréscimos, de minuto em minuto, por exemplo, deverâo ser negativos se **q < q**o, isto é, se as viagens médias foram superestimadas; e positivos, no caso contrário. Esse método permite um bom ajuste nessa área particularmente sensível da estimativa do padrâo de deslocamentos que é a proporçâo de viagens médias e longas.

4 Exemplos e resultados

As duas figuras que seguem apresentam alguns resultados ilustrativos do padrâo de mobilidade em Sâo Paulo. A Fig.3 permite uma comparação das distribuiçôes das viagens segundo o tempo de percurso por transporte individual, por um lado, e coletivo, por outro. Observa-se, como era de se esperar, uma concentraçâo das viagens em torno de uma média modal bem inferior para as primeiras (17'), contrastando com a parti-

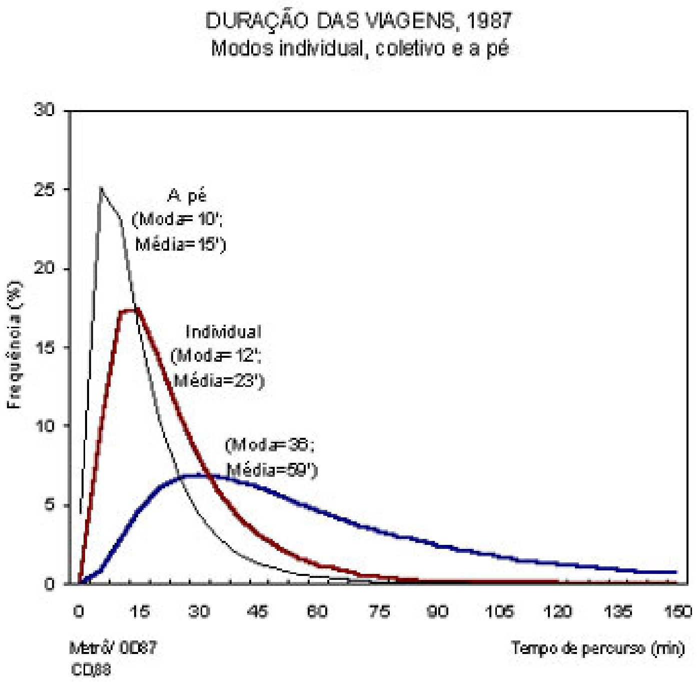


Figura 3

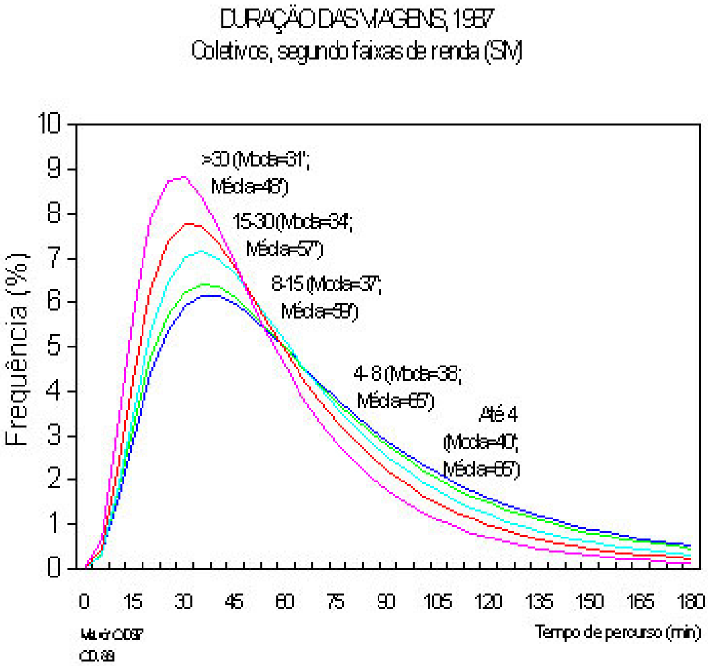


Figura 4

cipação bem superior das viagens mais longas e média modal igualmente superior (35')().

A Fig.4, por sua vez, ilustra a variação do padrão de mobilidade por transporte coletivo segundo a faixa de renda familiar da população, de acordo com cinco faixas 1-5 definidas como: até 3, 5, 10, 20 e mais de 20 salários mínimos, respectivamente. Aqui também, observa-se uma queda sensível da média modal (37, 34, 32, 30 e 21, respectivamente) segundo a ordem crescente de renda. Para clareza, está omissa na figura a curva correpondente à faixa de renda 3, inserida entre aquelas das faixas vizinhas 2 e 4, já bastante próximas.

5 Discussao

A calibração de uma funçãof(t) a partir da expressão (**2**) e para o nível de desagregação desejada é, como vimos, um processo bastante simples. Seus detalhes operacionais serão facilitadas ainda mais mediante o uso de planilhas de cálculo pré-programadas (tipo *Lotus* ou similar). A função distância .ajustada pode então ser utilizada para estimar a matriz O/D das viagens. Inicialmente calcula-se uma matriz de probabilidades relativas de oportunidades de viagens como em **(1)**:

$$\mathbf{v}_{ij} = f(\mathbf{t}_{ij}) \cdot \mathbf{P}_i \cdot \mathbf{E}_j \ , \quad \mathbf{(1)}$$

para a seguir obter cada linha **i** da matriz fazendo

$$\mathbf{v}_{ij} = \mathbf{v}_i \cdot \frac{\mathbf{v}_{ij}}{\Sigma_j \mathbf{v}_{ij}} \ , \quad \mathbf{(6)}$$

onde $\mathbf{v}_i$ é a produção de viagens na desagregaçâo relevante obtida mediante a expansâo direta da amostra, em cada zona**i**, **e** $\mathbf{v}_{ij}$, as viagens estimadas para cada zona **j**.

Observe-se, contudo, que mobilidade nâo é o mesmo que acessibilidade (ou atratividade) e assim, a mesma funçâo $f(t)$ nâo poderia ser utilizada para construir índices de atratividade ou de potencial. O que permite, aliás, a construçâo do modelo de deslocamentos acima é precisamente a distinçâo entre os conceitos de atratividade e de mobilidade. O modelo gravitacional, presa à idéia de uma fundamentaçâo "teórica" por analogia a alguma lei da física (no caso, à lei da gravidade), nâo pode admitir uma funçâo distância -- interpretada como o inverso de atratividade-- que nâo seja monotonicamente decrescente, o que constituiu uma dificuldade crônica e congênita para sua calibraçâo para simular -- fluxos de tráfego, isto é, *mobilidade*. Já o modelo aqui apresentado, livre daquela restriçâo, pode admitir a queda de frequência das viagens curtas e representa apenas isto: a probabilidade de uma viagem ocorrer em funçâo da distância associada a uma oportunidade de viagem -- razâo pela qual o chamamos, acima, de modelo probabilístico.