The Estimation of an Origin-Destination Matrix from Traffic Counts

M. G. H. BELL

University College London, London

In this paper, a model is described that will under certain circumstances yield the most likely origin-destination matrix which is consistent with measurements of link traffic volumes. Given knowledge of the proportionate usage of each link by the traffic between each zone pair, the model parameters scale each element of a prior estimate of the origin-destination matrix up or down so that the measurements of link traffic volumes are reproduced. The fitted values are shown by example to be invariant to the application of uniform scaling to the prior estimates. A Newton model fitting procedure is outlined. In reality, measurements of link traffic volumes are random variables and the proportionate usage of links by the traffic for each zone pair is not known with certainty. An expression is derived for the variances and covariances of the logarithms of the fitted values in terms of the variances and covariances of the measured link volumes, taking the proportionate usage of links as given. The expression permits the calculation of asymmetric confidence intervals for the elements of the fitted origin-destination matrix.

INTRODUCTION

There has recently been considerable interest in methods for estimating the most likely origin-destination (O-D) matrix (sometimes referred to as a trip table) from measurements of traffic volumes on the links in a network. Although it is possible to obtain O-D information directly by surveying trip makers (using, for example, household, workplace or roadside interviews), this tends to be both costly and labor intensive. By contrast, measurements of traffic volumes on links in a network can be made relatively inexpensively and, when automatic traffic counters are

Transportation Science
Vol. 17, No. 2, May 1983

0041-1655/83/1702-0198 $01.25

used, very little labor is required. Moreover, it is not necessary (as it is, for example, in the case of a roadside interview) to stop vehicles, so the associated risk of causing a disruption to traffic (possibly also resulting in a modification of its distribution) is avoided.

Some early approaches to the inference of an O-D matrix from measurements of link traffic volumes have been reviewed by WILLUMSEN.[1] JÖRNSTEN AND NGUYEN[2] have described an approach based on equilibrium traffic assignment which subsequently formed the basis of a program developed by the U.S. DEPARTMENT OF TRANSPORT.[3] VAN ZUYLEN AND WILLUMSEN[4] have developed closely related models for the estimation of an O-D matrix, Van Zuylen using information minimization and Willumsen entropy maximization. The associated problem of inferring turning movements within an isolated junction given measurements of traffic volumes on the approach roads has also received attention (BEIL,[5] CREMER AND KELLER,[6] HAUER ET AL.[7]). However, this problem is qualitatively different since only for an isolated junction is it generally true that flows entering and leaving the network can be measured by traffic counts.

A further model for the inference of an O-D matrix from link traffic volumes, originally suggested by VAN ZUYLEN[8] in order to overcome anomalies encountered with earlier models,[1,4,9] is given a probabilistic derivation. This new model, unlike the earlier models, generates an estimated O-D matrix that is always invariant to the application of uniform scaling to the prior estimates.

In common with the earlier models, a matrix of route choice proportions is required to be known. In practice, such a matrix may be hard to obtain. In the absence of direct observations, an assignment algorithm will have to be used. This problem has been discussed elsewhere.[1,4]

Making use of the matrix of route choice proportions, some properties of flow within a complex network, namely linear dependency and correlation, are examined. Linear dependencies are used to reduce the number of model parameters to be fitted. A Newton model fitting procedure is described, and some results concerned with the convergence of the procedure are given. An expression for the approximate variances and covariances of the logarithms of the fitted values (the estimated origin-destination movements) is derived assuming that route choice proportions are fixed but allowing link volumes to be random variables. The expression takes correlations between link volumes into account.

Following POTTS AND OLIVER,[10] the transport network is characterized by a graph of nodes connected by either directed or undirected links. Geographical areas which generate or attract traffic are represented by *zone centroids*, which are connected to the nodes of the graph by *centroid* *connectors.* Nodes which are not connected in this way to zone centroids are referred to as *intermediate nodes.* The level of resolution of the representation of the transport network can be adjusted to fit the available data. For example, when information about turning movements at junctions is available, these movements can each be represented by a link in the graph. It is therefore possible to associate with each movement within the transport network about which there is some information a particular link in the corresponding graph.

2. LINEAR DEPENDENCY IN COMPLEX NETWORKS

IN THIS PAPER, steady state traffic flows are considered, effectively removing the time dimension from the problem. By observing the propagation over the network of fluctuations in the levels of link flows, it is possible to make additional inferences about the nature of the distribution of traffic. Such inferences can be used to further constrain the set of feasible O-D matrices. For example, Cremer and Keller[6] suggest a method for estimating turning movements at an intersection by correlating variations in the flow of traffic on the links into and out of it, making allowance for the time required for vehicles to pass between entrance and exit points of measurement.

It is assumed here that Kirchoff's law applies within the network. In this context, Kirchoff's law asserts that traffic does not appear or disappear at nodes or in links. Under steady state conditions Kirchoff's law implies that the total flow into a link or node is equal to the total flow out. These conservation relationships reduce the degrees of freedom for link volumes, since some are linearly dependent on others.

Kirchoff's law has implications at two levels (see WILLUMSEN[11]). At the level of the node, conservation requires that there be one linearly dependent link volume for each intermediate node. At a higher level, conservation of flows on routes may lead to additional linear dependencies. For example, if two directed links lie on one and only one route, then Kirchoff's law requires that the flow on each be equal, even though the two links may not be connected to a common intermediate node.

Define

t_j = the quantity of traffic between a pair of zones designated by j (there being J zone pairs)

v_i = the volume of traffic on measured link i (there being I measured links); it is not assumed that all links in the network are measured

p_{ij} = the proportion of the traffic between zone pair j that uses measured link i.

Using the vector and matrix notation

$$\underline{t}^T = [t_1, \cdots, t_J]$$

$$\underline{v}^T = [v_1, \cdots, v_I]$$

$$\underline{\underline{P}} = \begin{bmatrix} p_{11}, & \cdots, & p_{1J} \\ \vdots & & \vdots \\ p_{I1}, & \cdots, & p_{IJ} \end{bmatrix}$$

the available network information can be expressed as the following set of simultaneous equations

$$\underline{v} = \underline{\underline{P}}\underline{t} \tag{1}$$

The number of linearly independent link volumes is identically equal to the rank of matrix $\underline{\underline{P}}$.

It is assumed throughout this paper that the matrix of route assignment proportions, $\underline{\underline{P}}$, is known. In practice, this may have to be obtained by making some assumption about route choice behavior. It should be noted, however, that it is not in general sufficient to assume that the network is user optimized (i.e. that Wardrop equilibrium obtains) since the equilibrium assignment model will not usually yield unique estimates for route choice proportions (see WILLUMSEN[11]).

Given a set of observed link volumes, $\underline{v}^*$, and the matrix of route assignment proportions, the set of feasible solutions for $\underline{t}$ is defined by relationship (1). There are a number of possibilities. First, there may be no feasible solution. When this is the case, at least one row of matrix $\underline{\underline{P}}$ is linearly dependent and the equations are said to be inconsistent (see HADLEY[12]). Although the inconsistency may be due to the violation of Kirchoff's law somewhere in the network, this is not necessarily the explanation; it may be the result of an incorrect hypothesized route assignment. Second, feasible solutions for $\underline{t}$ may exist but none of them may be strictly nonnegative (negative values for the elements of $\underline{t}$ have no meaning). This situation may arise if there is a divergence between actual and hypothesized route assignment. Third, while strictly nonnegative feasible solutions may exist, none of them may be strictly positive. This may reflect an underlying truth (namely that t_j does in fact equal zero for some j), but is more likely to be the result of a divergence between hypothesized and actual route assignment.

If any feasible solutions exist, then the rank of matrix $\underline{\underline{P}}$ is equal to the rank of the augmented matrix of the system of equations (the matrix $[\underline{\underline{P}}, \underline{v}^*]$ formed by appending vector $\underline{v}^*$, as an additional column, to matrix $\underline{\underline{P}}$) (see Willumsen[11]). When all the rows of matrix $\underline{\underline{P}}$ are linearly independent, the rank of $\underline{P}$ is equal to the rank of the augmented matrix and therefore at least one feasible solution exists.

The procedures described in subsequent sections require the identification of a set of linearly independent link volumes. This is equivalent to determining a set of rows of matrix $\underline{P}$ which may be regarded as being linearly dependent on the other rows. There are a number of suitable numerical methods for doing this, of which two (Gram-Schmidt orthogonalization and Gaussian elimination) are described here.

The use of Gram-Schmidt orthogonalization for the sequential search for linear dependency among the rows of matrix $\underline{P}$ has been independently suggested by the author and VAN ZUYLEN AND BRANSTON.[13]

Let $\underline{p}_i$ be the ith row of $\underline{P}$. On the rth step of the procedure

$$\underline{g}_r = \underline{p}_r - \textstyle\sum_{i=1}^{r-1} (\underline{g}_i^T \underline{p}_r)\underline{g}_i/(\underline{g}_i^T \underline{g}_i), \qquad r = 2, \cdots, n \tag{2}$$

where $\underline{g}_1$ is equal to $\underline{p}_1$. Vectors $\underline{g}_r$ and $\underline{g}_s$ ($r \neq s$) are orthogonal (i.e. $\underline{g}_r^T \underline{g}_s = 0$). Let a_{ir} be equal to $(\underline{g}_i^T \underline{p}_r)/(\underline{g}_i^T \underline{g}_i)$. Equation 2 can be written as

$$\underline{g}_r = \underline{p}_r - \textstyle\sum_{i=1}^{r-1} a_{ir}\underline{g}_i = \underline{p}_r - \sum_{i=1}^{r-1} a'_{ir}\underline{p}_i, \qquad r = 2, \cdots, n \tag{3}$$

where a'_{ir} are new coefficients formed from a_{jr} ($j < r$). If $\underline{g}_r = \underline{0}$, then the rth row of matrix $\underline{P}$ is a linear combination of the preceding rows (with coefficients $a'_{1r}, \cdots, a'_{r-1,r}$) and

$$v_r = \textstyle\sum_{i=1}^{r-1} a'_{ir} v_i. \tag{4}$$

In fact, $\underline{g}_r = \underline{0}$ is both a necessary and sufficient condition for linear dependency.

In the case of Gaussian elimination, matrix $\underline{P}$ is reduced by the successive elimination of the variables in (1) to a new matrix with zeros below the principal diagonal. If the rth row of this new matrix contains only zeros then the corresponding row of matrix $\underline{P}$ is linearly dependent on the preceding rows. As before, the reduction of a row to zero elements only is both a necessary and sufficient condition for linear dependency. Further information about the Gram-Schmidt orthogonalization process and Gaussian elimination may be found in Hadley[12] or ISAACSON AND KELLER.[14]

The result which follows, and which constitutes a key component of later proofs, asserts that the matrix $\underline{P}\underline{P}^T$ could equally well be analyzed for linear dependency. Typically, the number of zone pairs, J, greatly exceeds the number of measured links, I, so the matrix $\underline{P}\underline{P}^T$, which is $I \times I$, may be significantly smaller than matrix $\underline{P}$, which is $I \times J$.

RESULT 1. *The rth row of the matrix* $\underline{P}\underline{P}^T$ *is linearly dependent on the preceding rows if and only if the rth row of matrix* $\underline{P}$ *is linearly* *dependent on its preceding rows. Moreover, the linear combinations involved are the same.*

Proof. Let $\underline{P}$ have rows equal to $\underline{p}_1, \cdots, \underline{p}_I$ and $\underline{P}\underline{P}^T$ have rows equal to $\underline{z}_1, \cdots, \underline{z}_I$. Note that $\underline{z}_i^T = \underline{p}_i^T \underline{P}^T$.
(i) Necessity. If

$$\underline{p}_r = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{p}_i \tag{5}$$

where $\lambda_1, \cdots, \lambda_{r-1}$ are not all equal to zero, then

$$\underline{p}_r^T \underline{P}^T = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{p}_i^T \underline{P}^T \tag{6}$$

(by post-multiplying by $\underline{P}^T$). Thus

$$\underline{z}_r = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{z}_i \tag{7}$$

(ii) Sufficiency. If

$$\underline{z}_r = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{z}_i \tag{8}$$

where, as before, $\lambda_1, \cdots, \lambda_{r-1}$ are not all zero, then

$$\underline{p}_r^T \underline{P}^T = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{p}_i^T \underline{P}^T \tag{9}$$

or

$$\underline{u}^T \underline{P}^T = \underline{0}^T \tag{10}$$

where

$$\underline{u} = \underline{p}_r - \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{p}_i. \tag{11}$$

Vector $\underline{u}$ is both contained in the space spanned by $\underline{p}_1, \cdots, \underline{p}_r$ (see (11)) and orthogonal to $\underline{p}_1, \cdots, \underline{p}_r$ (see (10)). Therefore, $\underline{u}$ is equal to $\underline{0}$ and

$$\underline{p}_r = \textstyle\sum_{i=1}^{r-1} \lambda_i \underline{p}_i. \tag{12}$$

□

An immediate implication of this result is that the rank of $\underline{P}\underline{P}^T$ is equal to the rank of $\underline{P}$.

The matrix $\underline{P}\underline{P}^T$ has an interesting interpretation. Its (i, j)th element, equal to $\underline{p}_i^T \underline{P}_j$, has the value zero if there is no route which carries some traffic and uses both links i and j, and is positive otherwise. From (1), we have that

$$\underline{V}(\underline{v}) = \underline{P}\underline{V}(\underline{t})\underline{P}^T \tag{13}$$

where $\underline{V}(\underline{v})$ and $\underline{V}(\underline{t})$ are variance-covariance matrices for the elements of $\underline{v}$ and $\underline{t}$ respectively. If the elements of $\underline{t}$ (the trip table) are independent random variables with variance σ^2, then

$$\underline{\underline{V}}(\underline{v}) = \sigma^2 \underline{\underline{P}}\underline{\underline{P}}^T. \tag{14}$$

Thus the covariance between v_i and v_j is equal to $\sigma^2 \underline{p}_i^T \underline{p}_j$, or alternatively their correlation is equal to $(\underline{p}_i^T \underline{p}_j)/(|\underline{p}_i||\underline{p}_j|)$ where $|\underline{p}_k| = \sqrt{(\underline{p}_k^T \underline{p}_k)}$.

3. MODEL FORMULATION

DEFINE additional notation

$\tilde{t}_j$ = a prior estimate of t_j

q_j = the probability that a trip is between zone pair j given that a trip occurs.

Also let

$$\underline{\tilde{t}}^T = [\tilde{t}_1, \cdots, \tilde{t}_J]$$

$$\underline{\xi}^T = [\xi_1, \cdots, \xi_J] = [\ln t_1, \cdots, \ln t_J]$$

$$\underline{\tilde{\xi}}^T = [\tilde{\xi}_1, \cdots, \tilde{\xi}_J] = [\ln \tilde{t}_1, \cdots, \ln \tilde{t}_J]$$

$$\underline{q}^T = [q_1, \cdots, q_J].$$

A strictly positive solution for $\underline{t}$ in Equation 1, if it exists, is unlikely to be uniquely defined because in practice the number of unknowns, J, usually greatly exceeds the number of linearly independent equations. In this situation, a strictly convex, single-valued objective function, represented by $f(\underline{t})$, can be introduced in order to obtain a unique solution. The use of different forms of objective function in the constrained minimization problem

$$\mathrm{Min}_{\underline{t}} f(\underline{t}) \qquad \text{subject to } \underline{v} = \underline{\underline{P}}\underline{t} \tag{15}$$

yield a family of models, including those considered by Van Zuylen and Willumsen,[4] Willumsen and Van Vliet,[9] Beil,[5] and the model derived here.

From entropy maximizing considerations, Van Zuylen and Willumsen[4] obtained the following objective function

$$f(\underline{t}) = \textstyle\sum_j t_j(\ln(t_j/\tilde{t}_j) - 1). \tag{16}$$

The resulting model

$$t_j = \tilde{t}_j \Pi_i (\chi_i)^{p_{ij}}, \qquad j = 1, \cdots, J \tag{17}$$

has one parameter, χ_i, for each link volume which scales the prior estimate up or down so that the measured link volumes are reproduced. Willumsen and Van Vliet[9] suggest that (16) constitutes an appropriate measure of the difference between $\underline{t}$ and $\underline{\tilde{t}}$. If a constant term, $\sum_j \tilde{t}_j$, is added to (16), the resulting function is the sum of J strictly convex functions of one variable, t_j, each of which takes a value zero if $t_j = \tilde{t}_j$ and is positive otherwise. This interpretation is particularly attractive when the model is used to update a previous estimate of $\underset{\sim}{t}$ (perhaps obtained by a roadside, household or workplace interview) in the light of more recent measurements of traffic volumes.

Fitted values $\underset{\sim}{t}$ for model (17) may be obtained by solving the equations

$$v_i = \sum_j p_{ij}\tilde{t}_j \Pi_k (\chi_k)^{p_{kj}}, \qquad i = 1, \cdots, I \tag{18}$$

perhaps by a multiproportional adjustment algorithm, a special case of the more general balancing method studied by Bregman (LAMOND AND STEWART[15]), or by a Newton method. The multiproportional adjustment algorithm has been embodied in a program referred to as ME2 (Maximum Entropy Matrix Estimation) (WILLUMSEN[16]). It can be shown that these fitted values are not in general invariant to the application of uniform scaling to the prior estimates. This is a disadvantage when there has been a significant growth (or decline) in the total quantity of traffic between the time that the prior estimates were obtained and the time that the traffic counts were made.

Here a model is derived that maximizes the joint probability of observing $\underset{\sim}{t}$ subject to the constraints imposed by relationship (1), given the probabilities $\underset{\sim}{q}$ and the assumption that trips are multinomially distributed. If $f(\underset{\sim}{t})$ is this joint probability, then

$$f(\underset{\sim}{t}) = ((\textstyle\sum_j t_j)!/\Pi_j t_j!)\Pi_j (q_j)^{t_j} \tag{19}$$

and

$$F(\underset{\sim}{t}) = \ln f(\underset{\sim}{t}) = \ln(\textstyle\sum_j t_j)! - \sum_j \ln t_j! + \sum_j t_j \ln q_j. \tag{20}$$

Taking $d \ln x!/dx$ to be equal to $\ln x$, we obtain

$$(\partial F(\underset{\sim}{t})/\partial t_k) = \ln(\textstyle\sum_j t_j) - \ln t_k + \ln q_k, \qquad k = 1, \cdots, J. \tag{21}$$

The maximization of $F(\underset{\sim}{t})$ with respect to relationship (1) yields the following Lagrangian equation

$$L(\underset{\sim}{t}, \underset{\sim}{\lambda}) = F(\underset{\sim}{t}) + \underset{\sim}{\lambda}^T(\underset{\sim}{v} - \underset{=}{P}\underset{\sim}{t}) \tag{22}$$

where $\underset{\sim}{\lambda}$ is a vector of Lagrange multipliers. The necessary conditions for a maximum are

$$\partial L/\partial \underset{\sim}{t} = (\partial F(\underset{\sim}{t})/\partial \underset{\sim}{t}) - \underset{\sim}{\lambda}^T \underset{=}{P} = \underset{\sim}{0}^T \tag{23}$$

where $\partial L/\partial \underset{\sim}{t}$ is the row vector $[\partial L/\partial t_1, \cdots, \partial L/\partial t_J]$ and $\partial F/\partial \underset{\sim}{t}$ is similarly defined. The following result confirms that a maximum is achieved when (23) is satisfied.

RESULT 2. *The Hessian for $F(\underline{t})$, namely the matrix $[\partial^2 F/\partial t_i \partial t_j]$, is (subject to the constraint $\underline{v} = \underline{\underline{P}}\underline{t}$) negative definite for strictly positive $\underline{t}$.*

Proof. The Hessian of the objective function has the form

$$\partial^2 F/\partial t_i \partial t_j = \begin{cases} (1/\sum_k t_k) - (1/t_i) & \text{if } i = j \\ 1/\sum_k t_k & \text{otherwise.} \end{cases}$$

The quadratic form

$$\begin{aligned} \textstyle\sum_{ij} dt_i(\partial F/\partial t_i \partial t_j)\, dt_j &= \textstyle\sum_i dt_i \sum_j (\partial^2 F/\partial t_i \partial t_j)\, dt_j \\ &= \textstyle\sum_i dt_i((\sum_j dt_j/\sum_k t_k) - (dt_i/t_i)) \qquad (24) \\ &= (\textstyle\sum_i dt_i)^2/\sum_k t_k - \sum_i (dt_i^2/t_i) \end{aligned}$$

Let $\phi_i = dt_i/\sum_k dt_k$ and $\theta_i = t_i/\sum_k t_k$. For negative definiteness, we require

$$f(\underline{\theta}) = \textstyle\sum_i \phi_i(1 - (\phi_i/\theta_i)) < 0. \qquad (25)$$

To investigate $f(\underline{\theta})$ for maxima, consider the Lagrange equation

$$L(\underline{\theta}) = f(\underline{\theta}) - \omega(1 - \textstyle\sum_i \theta_i) \qquad (26)$$

where ω is a Lagrange multiplier. The necessary conditions for a maximum are

$$\partial L/\partial \theta_i = (\phi_i^2/\theta_i^2) - \omega = 0 \qquad \text{for all } i \qquad (27)$$

implying that $\phi_i = \sqrt{\omega}\theta_i$, and therefore that $\omega = 1$, since

$$\textstyle\sum_i \phi_i = \sum_i \theta_i = 1.$$

Furthermore, since $\theta_i > 0$ and

$$\partial^2 f/\partial \theta_i \partial \theta_j = \begin{cases} -\phi_i^2/\theta_i^3 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

the Hessian for $f(\underline{\theta})$ is negative definite. Thus, the sufficient condition for a maximum is satisfied. At this maximum, $\phi_i = \theta_i$, implying that

$$dt_i/\textstyle\sum_k dt_k = t_i/\sum_k t_k \qquad \text{for all } i. \qquad (28)$$

Moreover, the maximum value for $f(\underline{\theta})$ is zero. By differentiating the constraint equations, we obtain the relationship $d\underline{v} = \underline{\underline{P}}\, d\underline{t} = \underline{0}$. Thus, at the maximum of $f(\underline{\theta})$, $\underline{t} = \pi\, d\underline{t}$ (where $\pi \neq 0$ is a constant of proportionality) and therefore $\underline{\underline{P}}\underline{t} = \underline{0}$. This violates the constraints, since $\underline{v} \neq \underline{0}$. Hence $f(\underline{\theta}) < 0$ and consequently the Hessian for $F(\underline{t})$ is negative definite. □

From (21) and (23) we obtain

$$\ln t_k = \ln(\textstyle\sum_j t_j) + \ln q_k - \sum_i \lambda_i p_{ik}, \qquad k = 1, \cdots, J \qquad (29)$$

and thus

$$t_k = (\sum_j t_j) q_k \exp(- \sum_i \lambda_i p_{ik}), \qquad k = 1, \cdots, J. \tag{30}$$

In practice, the probabilities $\underline{q}$ are unknown. It is assumed here that they can be derived from the prior estimates as follows

$$q_k = \tilde{t}_k / \sum_j \tilde{t}_j, \qquad k = 1, \cdots, J. \tag{31}$$

This gives the following model

$$t_k = \tau \tilde{t}_k \exp(- \sum_i \lambda_i p_{ik}), \qquad k = 1, \cdots, J \tag{32}$$

where

$$\tau = \sum_j t_j / \sum_j \tilde{t}_j. \tag{33}$$

If $-\lambda_i$ is equal to $\ln \chi_i$ for all i, then the model becomes

$$t_k = \tau \tilde{t}_k \Pi_i (\chi_i)^{p_{ik}}, \qquad k = 1, \cdots, J. \tag{34}$$

This differs from the model given by (17) only in so far as a scale parameter τ is included. It is shown by example in Section 6 that the fitted values yielded by the new model are invariant to the application of uniform scaling to the prior estimates.

4. MODEL FITTING

IF ψ is equal to $\ln \tau$, the model as expressed in (32) may be rewritten as

$$t_k = \tilde{t}_k \exp(\psi - \sum_i \lambda_i p_{ik}), \qquad k = 1, \cdots, J. \tag{35}$$

There are now $I + 1$ parameters $(\psi, \lambda_1, \cdots, \lambda_I)$ to be determined by I equations

$$v_i = \sum_j p_{ij} t_j = \sum_j p_{ij} \tilde{t}_j \exp(\psi - \sum_k \lambda_k p_{kj}), \; i = 1, \cdots, I. \tag{36}$$

However, the number of linearly independent equations is equal to the rank of matrix $\underline{P}$ which is in turn equal to the number of linearly independent measured link volumes. Let L be the number of linearly independent measured link volumes.

Without loss of generality, let us suppose that the first L measured link volumes are linearly independent. Hence

$$v_i = \sum_j p_{ij} t_j, \qquad i = 1, \cdots, L \tag{37}$$

are linearly independent equations. The set of parameters can be correspondingly reduced, leaving $(\psi, \lambda_1, \cdots, \lambda_L)$ to be determined. There are now $L + 1$ parameters. The following additional relationship

$$\sum_j \tilde{t}_j = \tau^{-1} \sum_j t_j = \exp(-\psi) \sum_j t_j = \sum_j \tilde{t}_j \exp(- \sum_i \lambda_i p_{ij}) \tag{38}$$

is therefore required in order to obtain a unique solution.

Define further notation

$$\underline{y}^T = [\textstyle\sum_j \tilde{t}_j, v_1, \cdots, v_L] \tag{39}$$

and

$$\underline{\underline{R}} = \begin{bmatrix} \tau^{-1}, \cdots, \tau^{-1} \\ p_{11}, \cdots, p_{1J} \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ p_{L1}, \cdots, p_{LJ} \end{bmatrix}. \tag{40}$$

The $L + 1$ constraint equations can now be written as

$$\underline{y} = \underline{\underline{R}}\underline{t}. \tag{41}$$

If the rank of matrix $\underline{\underline{R}}$ is equal to that for matrix $\underline{\underline{P}}$, then these equations will not be linearly independent. Additionally define

$$\underline{\mu}^T = [\psi, -\lambda_1, \cdots, -\lambda_L] \tag{42}$$

and

$$\underline{\underline{S}} = \begin{bmatrix} 1, \cdots, 1 \\ p_{11}, \cdots, p_{1J} \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ \cdot \qquad \cdot \\ p_{L1}, \cdots, p_{LJ} \end{bmatrix} \tag{43}$$

thus enabling the model to be written as

$$\underline{t} = \tilde{\underline{\underline{D}}} \exp(\underline{\underline{S}}^T \underline{\mu}) \tag{44}$$

where $\tilde{\underline{\underline{D}}} = \mathrm{diag}(\underline{t}_1, \cdots, \underline{t}_J)$ and $\exp(\cdot)$ applies to each element of $\underline{\underline{S}}^T\underline{\mu}$. Alternatively, since the model is log-linear

$$\underline{\xi} = \tilde{\underline{\xi}} + \underline{\underline{S}}^T \underline{\mu} \tag{45}$$

There are a number of numerical solution procedures that may be used to derive values for the parameters, $\underline{\mu}$, so that the model estimates for $\underline{y}$ agree with predetermined values, $\underline{y}^*$. One such is the Newton procedure

$$\underline{\mu}^{(n+1)} = \underline{\mu}^{(n)} - (\underline{y}^{(n)})^{-1}(\underline{y}^{(n)} - \underline{y}^*) \tag{46}$$

where

$$\underline{y}^{(n)} = \underline{\underline{R}}\underline{t}^{(n)} = \underline{\underline{R}}\tilde{\underline{\underline{D}}} \exp(\underline{\underline{S}}^T \underline{\mu}^{(n)}) \tag{47}$$

and

$$\underline{\underline{J}}^{(n)}=\begin{bmatrix} \partial y_1^{(n)}/\partial\mu_1^{(n)}, & \cdots, & \partial y_1^{(n)}/\partial\mu_{L+1}^{(n)} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \partial y_{L+1}^{(n)}/\partial\mu_1^{(n)}, & \cdots, & \partial y_{L+1}^{(n)}/\partial\mu_{L+1}^{(n)} \end{bmatrix} \tag{48}$$

(superscript n denotes the iteration number). Moreover,

$$\partial y_i^{(n)}/\partial\mu_j^{(n)} = \begin{cases} 0 & \text{if } i=j=1 \text{ (see Equation 38)} \\ \sum_k (\partial y_i^{(n)}/\partial t_k^{(n)})\cdot(\partial t_k^{(n)}/\partial\mu_j^{(n)}) \\ = \begin{cases} \exp(-\mu_1^{(n)}) \sum_k p_{j-1,k} t_k^{(n)} \\ \quad = \exp(-\mu_1^{(n)}) v_{j-1}^{(n)} & \text{if } i=1 \text{ and } j>1 \\ \sum_k p_{i-1,k} t_k^{(n)} = v_{i-1}^{(n)} & \text{if } i>1 \text{ and } j=1 \\ \sum_k p_{i-1,k} p_{j-1,k} t_k^{(n)} & \text{if } i>1 \text{ and } j>1 \end{cases} \end{cases}$$

Thus

$$\underline{\underline{J}}^{(n)}=\left[\begin{array}{c|ccc} 0 & \exp(-\mu_1^{(n)})v_1^{(n)}, & \cdots, & \exp(-\mu_1^{(n)})v_L^{(n)} \\ \hline v_1^{(n)} & \sum_k p_{1k}^2 t_k^{(n)}, & \cdots, & \sum_k p_{1k}p_{Lk}t_k^{(n)} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ v_L^{(n)} & \sum_k p_{Lk}p_{1k}t_k^{(n)}, & \cdots, & \sum_k p_{Lk}^2 t_k^{(n)} \end{array}\right] \tag{49}$$

$$= \left[\begin{array}{c|c} 0 & \exp(-\mu_1^{(n)})\underline{v}^{(n)T} \\ \hline \underline{v}^{(n)} & \underline{\underline{P}}\,\underline{\underline{D}}^{(n)}\underline{\underline{P}}^T \end{array}\right].$$

The following two results prove that matrix $\underline{\underline{J}}^{(n)}$ is nonsingular and that therefore $\underline{\underline{J}}^{(n)-1}$ is determinate. In order for the Newton procedure to converge to the solution $\underline{\mu}^*$ in a neighborhood of it, it is sufficient that $\underline{\underline{J}}^{(n)}$ be continuous and nonsingular (see ORTEGA AND RHEINBOLDT[17]).

RESULT 3. *The matrix $\underline{\underline{P}}\,\underline{\underline{D}}^{(n)}\underline{\underline{P}}^T$ is positive definite for strictly positive prior estimates $\underline{\tilde{t}}$.*

Proof. It can be verified from (44) that given positive $\underline{\tilde{t}}$ the model yields positive fitted values $\underline{t}$. Thus $\underline{\underline{D}}^{(n)}$, which is equal to $\text{diag}(t_i^{(n)}, \ldots, t_J^{(n)})$, is nonsingular. Matrix $\underline{P}\underline{D}^{(n)}\underline{P}^T$ may be reformulated as $\underline{Q}\underline{Q}^T$ where $\underline{Q}$ is equal to $\underline{P}\sqrt{\underline{D}^{(n)}}$ and $\sqrt{\underline{D}^{(n)}}$ is equal to diag$(\sqrt{t_1^{(n)}}, \ldots, \sqrt{t_J^{(n)}})$. Since $\underline{D}^{(n)}$ is nonsingular, $\sqrt{\underline{D}^{(n)}}$ is also nonsingular and therefore the rank of $\underline{Q}$ is equal to the rank of $\underline{P}$. However, Result 1 implies that the rank of $\underline{Q}\underline{Q}^T$ is equal to the rank of $\underline{Q}$, so the rank of $\underline{P}\underline{D}^{(n)}\underline{P}^T$ is equal to the rank of $\underline{P}$. Thus, since the rank of $\underline{P}$ is equal to the number of its rows, L, the matrix $\underline{P}\underline{D}^{(n)}\underline{P}^T$, which is $L \times L$, must be nonsingular. Since it is expressible in the form $\underline{Q}\underline{Q}^T$, it must also be positive definite. □

RESULT 4. *The Jacobian* $\underline{J}^{(n)}$, *given by* (49), *is nonsingular.*

Proof. Matrix $\underline{P}\underline{D}^{(n)}\underline{P}^T$, which is $L \times L$, has rank L (Result 3). Hence columns 2 to $L + 1$ of $\underline{J}^{(n)}$ are linearly independent. Therefore, if $\underline{J}^{(n)}$ is singular, then

$$\begin{bmatrix} 0 \\ \underline{v}^{(n)} \end{bmatrix} = \begin{bmatrix} \exp(-\mu_1^{(n)})\underline{v}^{(n)T} \\ \underline{P}\underline{D}^{(n)}\underline{P}^T \end{bmatrix} \underline{\zeta} \tag{50}$$

where the elements of $\underline{\zeta}$ are not all zero. Thus

$$\left.\begin{aligned} \exp(-\mu_1^{(n)})\underline{v}^{(n)T}\underline{\zeta} &= 0 \\ \underline{P}\underline{D}^{(n)}\underline{P}^T\underline{\zeta} &= \underline{v}^{(n)} \end{aligned}\right\} \tag{51}$$

Hence, since $\exp(-\mu_1^{(n)}) > 0$

$$\underline{\zeta}^T\underline{P}\underline{D}^{(n)}\underline{P}^T\underline{\zeta} = \underline{\zeta}^T\underline{v}^{(n)} = 0 \tag{52}$$

which contradicts the result that $\underline{P}\underline{D}^{(n)}\underline{P}^T$ is positive definite (Result 3). Therefore $\underline{J}^{(n)}$ must be nonsingular. □

Results 3 and 4 imply that the number of zone pairs, J, should be greater than or equal to L, the number of rows of matrix $\underline{P}$. If not, the number of columns of $\underline{P}$ must be less than L implying that the rank of $\underline{P}$, and therefore $\underline{P}\underline{D}^{(n)}\underline{P}^T$, is less than L. Since $\underline{P}\underline{D}^{(n)}\underline{P}^T$ is no longer positive definite, $\underline{J}^{(n)}$ is no longer necessarily nonsingular.

5. APPROXIMATE VARIANCES AND COVARIANCES FOR THE FITTED VALUES

AN IMPORTANT consideration for the practitioner is the robustness of the fitted values, $\hat{\underline{t}}$ ($\hat{}$ refers to estimated quantities), given that the measurements for $\underline{v}$ are random variables and that the set of assignment proportions, $\underline{P}$, are not known with certainty. As a first approximation to the problem, it is assumed here that the assignment proportions are known with certainty. An expression is derived for the variances and covariances of the logs of the fitted values, $\underline{\xi}$, in terms of the variances and covariances for the mean measurments of link traffic volumes, $\underline{v}^*$ (* refers to observed values).

Variances and covariances are obtained for $\hat{\underline{\xi}}$, rather than $\hat{\underline{t}}$, because it is conventional to assume that the elements of the former, rather than the latter, are approximately normally distributed. Standard errors for $\hat{\underline{\xi}}$ permit the construction of asymmetric confidence intervals for $\hat{\underline{t}}$ which exclude the possibility of negative (or zero) values. Additional advantages of this approach are that the standard errors for $\hat{\underline{\xi}}$ may be interpreted as approximate coefficients of variation for $\hat{\underline{t}}$, that no specific error structure hypothesis is assumed, and that correlations between traffic volumes are taken into account.

The expression is based on a truncated Taylor series approximation. For $\underline{x}$ close to $\underline{\eta}$

$$\underline{f}(\underline{x}) \simeq \underline{f}(\underline{\eta}) + [\partial \underline{f}/\partial \underline{x}]_{\underline{x}=\underline{\eta}}(\underline{x} - \underline{\eta}). \tag{53}$$

Let $E\{\underline{x}\} = \underline{\eta}$, where $E\{\underline{x}\}$ denotes the expectation of $\underline{x}$, and let matrix

$$\underline{\underline{Z}} = [\partial \underline{f}/\partial \underline{x}]_{\underline{x}=\underline{\eta}} \tag{54}$$

be the Jacobian of the transformation. Hence $E\{\underline{f}(\underline{x})\} \simeq \underline{f}(\underline{\eta})$ and the variance-covariance matrix for $\underline{f}(\underline{x})$ is given by the following approximation

$$\begin{aligned} \underline{\underline{V}}(\underline{f}(\underline{x})) &= E\{(\underline{f}(\underline{x}) - \underline{f}(\underline{\eta}))(\underline{f}(\underline{x}) - \underline{f}(\underline{\eta}))^T\} \\ &\simeq E\{\underline{\underline{Z}}(\underline{x} - \underline{\eta})(\underline{x} - \underline{\eta})^T \underline{\underline{Z}}^T\} = \underline{\underline{Z}}\,\underline{\underline{V}}(\underline{x})\underline{\underline{Z}}^T. \end{aligned} \tag{55}$$

Two stages are required in order to obtain a variance-covariance matrix for $\hat{\underline{\xi}}$ from a variance-covariance matrix for $\underline{y}^*$. First a variance-covariance matrix is obtained for $\hat{\underline{\mu}}$, then from this a variance-covariance matrix is derived for $\hat{\underline{\xi}}$. Since the first element of vector $\underline{y}^*$, equal to $\sum_j \tilde{t}_j$, is invariant, the first row and column of the variance-covariance matrix for $\underline{y}^*$ have zero elements only.

Using approximation (55)

$$\underline{\underline{V}}(\underline{y}^*) \simeq \underline{\underline{J}}\,\underline{\underline{V}}(\hat{\underline{\mu}})\underline{\underline{J}}^T \tag{56}$$

where the Jacobian matrix $\underline{\underline{J}}$ is as defined in Equation 52 when $\underline{\mu}^{(n)}$ is equal to $\hat{\underline{\mu}}$. Since $\underline{\underline{J}}$ is nonsingular (Result 4)

$$\underline{\underline{V}}(\hat{\underline{\mu}}) \simeq \underline{\underline{J}}^{-1}\underline{\underline{V}}(\underline{y}^*(\underline{\underline{J}}^T)^{-1} = \underline{\underline{J}}^{-1}\underline{\underline{V}}(\underline{y}^*)(\underline{\underline{J}}^{-1})^T. \tag{57}$$

Following Equation 45, the fitted model can be written as

$$\hat{\underline{\xi}} = \tilde{\underline{\xi}} + \underline{\underline{S}}^T\hat{\underline{\mu}}. \tag{58}$$

Thus

$$\partial\hat{\underline{\xi}}/\partial\hat{\underline{\mu}} = \underline{\underline{S}}^T. \tag{59}$$

Hence

$$\underline{\underline{V}}(\hat{\underline{\xi}}) \simeq \underline{\underline{S}}^T\underline{\underline{V}}(\hat{\underline{\mu}})\underline{\underline{S}} \simeq \underline{\underline{S}}^T\underline{\underline{J}}^{-1}\underline{\underline{V}}(\underline{y}^*)(\underline{\underline{J}}^{-1})^T\underline{\underline{S}}. \tag{60}$$

Standard errors for $\hat{\underline{\xi}}$ can be obtained by taking the square root of the elements on the principal diagonal of matrix $\underline{\underline{V}}(\hat{\underline{\xi}})$. Since maximum likelihood estimators are normally distributed for large samples (see, for example, EDWARDS[18]), confidence intervals for $\hat{\underline{\xi}}$ may be obtained from the unit normal distribution.

In practice, a variance-covariance matrix for link traffic volumes may be obtained by making repeated measurements on each link. For this purpose, the survey period can be divided into intervals. In each interval, the traffic can be counted on a subset of links with linearly independent traffic volumes. The traffic count variances for each link, and the covariances for each pair of links, can then be calculated.

The extent of the correlation between link volumes will depend on the size of the chosen interval of measurement. Since trips are not instantaneous, the correlations between link traffic volumes are in reality lagged. It is therefore desirable to select as long an interval of measurement as possible so as to minimize the significance of the lags.

It is assumed here that correlations between link volumes are instantaneous. Let x_{ij} be the count for the jth interval on link i expressed as a deviation from the mean count for link i.

Define

$$\underline{\underline{X}} = \begin{bmatrix} 0, & \cdots, 0 \\ x_{11}, & \cdots, x_{1N} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_{L1}, & \cdots, x_{LN} \end{bmatrix} \tag{61}$$

where L is the number of linearly independent measured links and N is the number of measurements made for each link. Thus

$$\underline{\underline{V}}(\underline{y}^*) = (1/N(N-1))(\underline{\underline{X}}\underline{\underline{X}}^T). \tag{62}$$

Substituting into approximation (60), and treating the approximation as an equality, we obtain

$$\underline{\underline{V}}(\hat{\underline{\xi}}) = \underline{\underline{S}}^T\underline{\underline{J}}^{-1}(1/N(N-1))(\underline{\underline{X}}\underline{\underline{X}}^T)(\underline{\underline{J}}^{-1})^T\underline{\underline{S}} \tag{63}$$

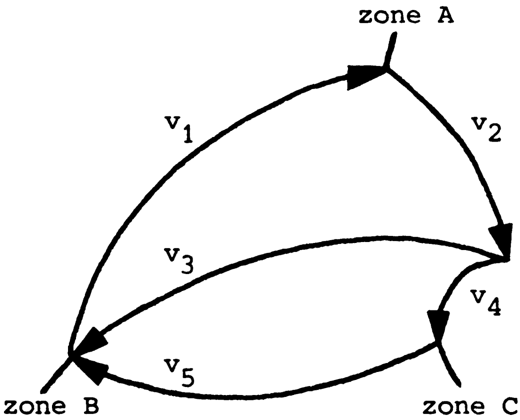

Fig. 1. Example network.

TABLE I
Matrix of Assignment Proportions

Link	Zone Pair					
	AB	*AC*	*BC*	*CB*	*CA*	*BA*
1	0	0	1	0	1	1
2	1	1	1	0	0	0
3	0.7	0	0	0	0	0
4	0.3	1	1	0	0	0
5	0.3	0	0	1	1	0

6. AN EXAMPLE

A COMPUTER program has been developed by the author to implement the methods described in the preceding sections. In this section, the results yielded by a simple network, depicted in Figure 1, are considered. There are six zone pairs, five links and one intermediate node. The assumed matrix of assignment proportions, which is not based on an all-or-nothing assignment (as evidenced by the existence of some noninteger values), is shown in Table I.

Hypothetical counts and their mean values are shown in Table II, and the corresponding variances and covariances in Table III. Taken by columns, the values in Table II conform to Kirchoff's law. Since the mean values also conform to Kirchoff's law, the system of constraint equations is consistent.

In this case, matrix $\underline{P}$ has rank 4, the linearly dependent row being due to flow conservation at the intermediate node. Either v_2, v_3 or v_4 may be treated as being linearly dependent. The Gaussian elimination method

TABLE II

Hypothetical Counts and Their Mean Values

Link	Measurement 1	2	3	4	5	Mean Value
1	26	13	23	13	21	19.2
2	27	21	17	20	19	20.8
3	14	13	10	11	6	10.8
4	13	8	7	9	13	10.0
5	12	13	14	11	15	13.0

TABLE III

Variances and Covariances for Hypothetical Counts

Link	Link 1	2	3	4	5
1	35.0				
2	7.0	14.0			
3	−1.2	7.9	9.7		
4	8.3	6.3	−1.8	8.0	
5	3.3	−3.0	−3.5	5.0	2.5

for detecting linear dependency is preferred in practice to the Gram-Schmidt orthogonalization process because, by applying the procedure to the augmented matrix of the system, it is simple to identify linearly dependent equations and simultaneously to check their consistency. The Gaussian elimination subroutine of the computer program suggests that link 4 is linearly dependent.

Taking uniform prior estimates equal to one for all zone pairs, the parameter estimates are

$$\hat{\mu}_1 = \hat{\psi} = 1.89$$

$$\hat{\mu}_2 = \hat{\lambda}_1 = 0.48$$

$$\hat{\mu}_3 = \hat{\lambda}_2 = -1.17$$

$$\hat{\mu}_4 = \hat{\lambda}_3 = 3.19$$

$$\hat{\mu}_5 = \hat{\lambda}_4 = -0.73.$$

The fitted values and their 95% confidence intervals are:

Origin	*Destination*	*Estimate*	*95% confidence region*
A	*B*	15.43	11.98 to 19.87
A	*C*	2.06	1.13 to 3.75
B	*C*	3.32	1.94 to 5.67
C	*B*	3.20	2.24 to 4.59
C	*A*	5.17	3.93 to 6.79
B	*A*	10.72	7.37 to 15.58.

If the prior estimates are multiplied by a factor of 10, then the parameter estimates are

$$\hat{\mu}_1 = -0.41$$
$$\hat{\mu}_2 = 0.48$$
$$\hat{\mu}_3 = -1.17$$
$$\hat{\mu}_4 = 3.19$$
$$\hat{\mu}_5 = -0.73.$$

Parameter $\hat{\mu}_1$ adjusts to allow for the change in the scaling applied to the prior estimates. The fitted values and their confidence intervals remain unchanged.

Five of the fitted values are sensitive to a change to nonuniform prior estimates. When the prior estimate for the movement from zone B to zone A is twice the prior estimates for all other zone pairs (suggesting that a trip from B to A is twice as likely as a trip between any other pair of zones, given that a trip occurs) the fitted values and their 95% confidence intervals are

Origin	*Destination*	*Estimate*	*95% confidence region*
A	B	15.43	11.98 to 19.87
A	C	2.64	1.49 to 4.69
B	C	2.73	1.59 to 4.70
C	B	4.12	2.99 to 5.68
C	A	4.25	3.21 to 5.64
B	A	12.22	8.76 to 17.03

Since (as Table I shows) the movement from A to B is determined by the movement on link 3, the corresponding fitted value and its confidence interval is insensitive to changes in the pattern of the prior estimates.

The variance-covariance matrix for the logarithms of the fitted values, obtained when the prior estimates are uniformly equal to one, is given in Table IV. It is clear from (63) that this matrix is symmetric, so only the lower triangular portion is given.

Since the movement from A to B is equal to 1.43 times the flow on link 3 (see Table I), the variance of the corresponding fitted value should be equal to 1.43 squared times the variance of the mean flow on link 3, namely 3.96. However, the program supplies an approximate variance for the logarithm of the fitted value (see Table IV). Since the standard deviation of the logarithm of the fitted value is approximately equal to the coefficient of variation for the fitted value itself, the program indicates a variance approximately equal to 4.05, which is of the correct order of magnitude.

TABLE IV

Variances and Covariances for the Logarithms of the Fitted Values

Zone Pair	Zone Pair					
	AB	*AC*	*BC*	*CB*	*CA*	*BA*
AB	0.017					
AC	−0.025	0.094				
BC	−0.018	0.076	0.075			
CB	−0.021	0.035	0.019	0.034		
CA	−0.014	0.016	0.018	0.018	0.019	
BA	0.010	−0.016	0.008	−0.021	0.003	0.036

7. DISCUSSION

A MODEL is developed in this paper that yields the most probable set of O-D movements that are consistent with a set of link flows if the probabilities obtained from the prior estimates are correct. In practice, the prior estimates might be based on an old survey that may not correctly reflect current probabilities. If this is so, or if there is no basis on which to obtain prior estimates (in which case a uniform value has to be adopted), it seems unlikely that the model will yield the most probable set of O-D movements consistent with link flows. As demonstrated in Section 6, at least some of the fitted values may be sensitive to changes in the prior estimates.

In practice, traffic counts are likely to be available for a set of links which include some whose volumes are linearly dependent on the volumes of the others within the set. When inconsistency arises, it is possible to remove it by adjusting the observations using a maximum likelihood criteria (see Van Zuylen and Willumsen[4] and Van Zuylen and Branston[13]), thus incorporating the additional information. However, it would in principle be preferable to fit the model directly to the traffic counts.

ACKNOWLEDGMENTS

THE AUTHOR is grateful to Dirck Van Vliet, Richard Allsop, Stephen Gallivan, David Holden, Henk Van Zuylen, and especially to Luis Willumsen, for helpful discussions at various stages during the development of the work, and wishes to express appreciation for the financial support provided by the Science and Engineering Research Council.

REFERENCES

1. L. G. WILLUMSEN, "Simplified Transport Models Based on Traffic Counts," *Transportation* **10,** 257–278 (1981).
2. K. O. JÖRNSTEN AND S. NGUYEN, "On the Estimation of a Trip Matrix from Network Data," Centre de Recherche sur les Transports Publication 153, University of Montreal, 1979.

3. U.S. DEPARTMENT OF TRANSPORT, "Estimation of an Origin-Destination Trip Table Based on Observed Link Volumes and Turning Movements," Federal Highway Administration, Offices of Research and Development, Traffic Systems Division, 1980.
4. H. J. VAN ZUYLEN AND L. G. WILLUMSEN, "The Most Likely Trip Matrix Estimated from Traffic Counts," *Trans. Res.* **14B,** 281–293 (1980).
5. D. BEIL, "Matrixmodelle zur Ermittlung von Teilströmen aus Querschnittsmessungen," thesis for Dr.-Ing., Fakultät für Bauingenieur und Vermessungswesen der Universität Karlsruhe, 1979.
6. M. CREMER AND H. KELLER, "Dynamic Identification of Flows from Traffic Counts at Complex Intersections," Proceedings 8th International Symposium on Transportation and Traffic Theory, Toronto University, June 1981.
7. E. HAUER, E. PAGITSAS AND B. T. SHIN, "Estimating Turning Flows from Automatic Counts," *Trans. Res. Rec.* **759,** 1–7 (1981).
8. H. J. VAN ZUYLEN, "Some Improvements in the Estimation of an O-D Matrix from Traffic Counts," Proceedings 8th International Symposium on Transportation and Traffic Theory, Toronto University, June 1981.
9. L. G. WILLUMSEN AND D. VAN VLIET, "Validation of the ME2 Model for Estimating Trip Matrices from Traffic Counts," Proceedings 8th International Symposium on Transportation and Traffic Theory, Toronto University, June 1981.
10. R. B. POTTS AND R. M. OLIVER, *Flows in Transportation Networks*, Academic Press, New York, 1972.
11. L. G. WILLUMSEN, "An Entropy Maximising Model for Estimating Trip Matrices from Traffic Counts," thesis for Ph.D., University of Leeds, 1981.
12. G. HADLEY, *Linear Algebra*, Addison-Wesley, Reading, Mass., 1979.
13. H. J. VAN ZUYLEN AND D. M. BRANSTON, "Consistent Link Flow Estimation from Counts," *Transportation Research* (forthcoming).
14. E. ISAACSON AND H. B. KELLER, *Analysis of Numerical Methods*, John Wiley & Sons, New York, 1966.
15. B. LAMOND AND N. F. STEWART, "Bregman's Balancing Method," *Trans. Res.* **15B,** 239–248 (1981).
16. L. G. WILLUMSEN, "Estimating a Trip Matrix from Traffic Counts, Empirical Tests," 13th Annual Conference of the Universities Transport Study Group, Leeds University, January 1981 (unpublished).
17. J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Non-linear Equations in Several Variables*, Academic Press, New York, 1970.
18. A. W. F. EDWARDS, *Likelihood*, Cambridge University Press, New Rochelle, N.Y., 1972.