3.0 Overview of Sub Problems

3.1 Introduction

A common class of traffic engineering and transportation planning problems revolves around the estimation of an Origin-Destination matrix for a network. This class of problems can take on many different formats, depending upon what is considered as known a priori, and what assumptions are made subsequently to derive the missing parameters. This Chapter attempts to provide a comprehensive overview of this entire class of problems and emphasizes both the various problem formulations and numerical solution approaches that can be considered. The objective is to provide this overview to potential users of the available techniques in such a manner as to permit them to make intelligent decisions as to which technique to utilize when, and to appreciate the implications of the approximations that are commonly made.

The first categorization, of the available estimation techniques, relates to whether the O-D's to be estimated are static, and apply to only one observation time period, or whether estimates are required for a series of linked dynamic time periods. This thesis will focus primarily on the static or single time period problem. The next breakdown relates to whether the estimation is based on input information about the magnitude of trip ends only, or whether input information is also available on additional links along the rest of the route of each trip. The former problem is commonly referred to as the trip distribution problem in demand forecasting, while the latter problem is commonly referred to as the synthetic O-D generation problem. The findings of the thesis apply to both problems, but will focus primarily on the latter synthetic O-D generation problem. Specifically, the former is simply viewed as a simpler subset of the latter problem.

Within the overall static synthetic O-D generation problem, there are two main flavors. The first exists when the routes, that vehicles are taking through the network, are considered as being known a priori. The second arises when these routes need to be estimated concurrently while the O-D demand is being estimated. A priori knowledge of routes can arise automatically, implicitly when there is only one feasible route between each O-D pair, or when observed traffic volumes are only provided for the zone connectors at the origins and destinations in the network. The first condition is common when O-D's need to be estimated for a single intersection or arterial, or a single interchange or freeway. The second condition is the default for any trip distribution analysis. Here we will focus on situations where the routes are known a priori, but it should be noted that a solution to the more general problem does exist. Furthermore, this more general solution involves iterative use of the solution approach for conditions where routes are considered as known a priori.

Within the static synthetic O-D generation problem, for scenarios where routes are known a priori (or are assumed to be known a priori), there exist two sub-problems. The first of these relates to situations where flow continuity exists at each node in the network, and multiple O-D matrices can be shown to match these observed flows exactly. In this case, the most likely of these multiple possible O-D matrices needs to be identified. The second sub-problem relates to situations where flow continuity does not exist at either the node level or at the network level. In other words, the observed traffic flows are inconsistent such that no matrix exists that will match the observed flows exactly. In this case, a new set of complementary link flows needs to be identified that permits flow continuity to be achieved, while introducing the least changes from the observed flows. Once such a complementary set of flows has been identified, the maximum likelihood problem can be solved as before using techniques that strictly require the input data to exhibit link flow continuity.

The static synthetic O-D generation problem, for scenarios where flow continuity does exit, can be formulated in two different ways. The first of these considers that the fundamental unit of measure is the individual trip, while the second considers that the fundamental unit of measure is the observation of a single vehicle on a particular link. The availability of a seed or target O-D matrix is implicit in the latter formulation, but can be dropped in the former formulation. However, it will be shown later that only when a seed matrix is properly included in the former formulation is it guaranteed to yield consistent results with the latter formulation. In other words, the absence of a seed matrix in the trip based formulation can be shown to yield inconsistent results, at least for networks in which the multiple possible solutions involve a different number of total network trips.

An additional and related attribute, of the trip-based formulation of maximum likelihood, is the presence of a term in the objective function that is based on the total number of trips in the network. This term, referred to as T, is often dropped in some approximations. However, it can be shown that dropping this term can yield solutions that represent only a very poor approximation to the true solution. In contrast, approximations that involve the use of Stirling's approximation, for representing the logarithm of factorials, are shown to yield consistently very good approximations. This finding is critical, as use of Stirling's approximation is critical to being able to compute the derivatives that are needed to numerically solve the problem (it is difficult to take derivatives of terms that include factorials).

Once Stirling's common approximation has been made, further approximations have been made in the past by Willumsen et al. These eventually permit the synthetic O-D generation problem to be reduced to a linear regression problem. While these approximations provide for solutions that sometimes are easier to compute, these solutions are also shown to considerably degrade the accuracy of the final solution (at least, when compared to the formulation without such an approximation).

The following sections will provide a more detailed discussion of the nature of each of the above facets of the O-D estimation problem, and will describe their implications for the accuracy or validity of the final O-D estimate.

3.2 Trip Distribution vs. Synthetic O-D Generation

Consider the network illustrated in Figure 3.1, which illustrates a set up in which there are 4 zones (A, B, C and D), and 16 links. Links 1, 3, 5 and 7 represent zone connectors leaving each zone, while links 2, 4, 6 and 8 represent zone connectors entering each zone. Furthermore, links 11, 13, 15 and 17 represent links that provide for clock-wise flow around the network, while links 12, 14, 16 and 18 provide for counter clockwise flow.

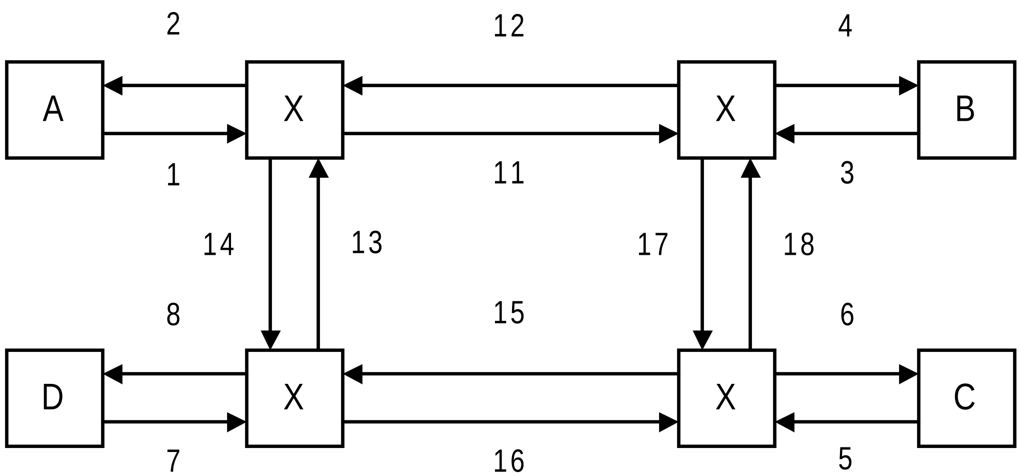

Figure 3.1: 16-Link Network

If the above problem were to be viewed as a trip distribution problem, traffic flows on only links 1 to 8 would be available. Specifically, links 1, 3, 5 and 7 would represent zonal trip productions, while links 2, 4, 6 and 8 would represent zonal attractions. In other words, within the normal trip distribution process, information on links 11 to 18 would not be considered explicitly. The absence, of any consideration of any en-route link flows also makes it unimportant to know the routing of traffic from, say A to D. In contrast, knowledge of the routes is critical during the synthetic O-D generation process, as the relevance of link counts to a particular O-D pair is highly dependent on the fraction of vehicles from that O-D pair that cross that link.

The complementary nature of the synthetic O-D generation process and the trip distribution process is best illustrated in Table 3.1. Note that it is assumed that the zone connectors are dummy links, rather than actual streets or groups of streets. Therefore they are not included in the synthetic O-D generation process as inputs. In contrast, if these zones were well defined subdivisions, business or industrial parks, or shopping centers, the zone connectors could represent actual or physical roads, and could therefore be counted, and included in the synthetic O-D generation process.

Table 3.1: Complementary Nature of Synthetic O-D Process and Trip Distribution

Link No.	Link Type	Complete Link Information	Information Used in Trip Distribution	Information Used in Synthetic O-D Generation
1	Zone Connector	100	100	-
2	Zone Connector	200	200	-
3	Zone Connector	200	200	-
4	Zone Connector	200	200	-
5	Zone Connector	300	300	-
6	Zone Connector	300	300	-
7	Zone Connector	400	400	-
8	Zone Connector	300	300	-
11	En-route Link	35	-	35
12	En-route Link	61	-	61
13	En-route Link	139	-	139
14	En-route Link	65	-	65
15	En-route Link	269	-	269
16	En-route Link	296	-	296
17	En-route Link	80	-	80
18	En-route Link	165	-	165

Similarly, it is considered that for the trip distribution process, the trip production and attraction rates are not measured using road counts, but are estimated from land use based on trip generation equations. Such forecasts are typically for 24 hours, or at best for peak periods. However, at present, trip-end estimates are not commonly available by 15-minute period, making the estimation of 15-minute O-D matrices using trip distribution methods virtually impossible. Of course, some planners scale 24 hour or peak period matrices, to produce 15-minute matrices, using traffic flow counts. However, if this scaling is performed ad hoc, a better approach would be to treat the trip distribution matrix as a seed matrix, and then systematically utilize the observed traffic flow counts in a synthetic O-D generation process. Table 3.2 clearly illustrates the relationship between trip distribution and synthetic O-D generation, for the 16-link network.

Table 3.2: Comparison of Trip Distribution vs. Synthetic O-D Generation

Trip Distribution				
	A	B	C	D
A	0.00	34.95	11.71	53.34
B	60.63	0.00	79.77	59.60
C	22.93	90.04	0.00	187.04
D	116.45	75.00	208.54	0.00
Synthetic O-D Generation				
	A	B	C	D
A	0.00	34.75	13.76	57.15
B	59.78	0.00	82.54	54.89
C	25.17	92.71	0.00	184.89
D	118.74	70.75	207.26	0.00

3.3 Multiple Solutions

Both the synthetic O-D generation process and the trip distribution process are subject to an information under-specification limitation. For example, if one attempts to estimate an O-D matrix for a 100-zone network with observed counts on, say 1000 links, one has many more unknowns to solve for than there are constraints. In the case of the trip distribution process, there are 100x100 O-D cells to be estimated, and only 2x100 trip end constraints. In the case of the synthetic O-D generation process, there are again 100x100 O-D cells to estimate, and only 1000 link constraints. Given the possibility of multiple solutions, both the trip generation process and the synthetic O-D generation process invoke additional considerations to select a preferred matrix from among the multiple solutions, as indicated next.

In the case of the synthetic O-D generation, the desire is to select from among all of the possible solutions, the most likely O-D matrix. This approach requires one to define a measure of the likelihood of each matrix. In general, there are two approaches to establish the likelihood of a matrix. One of them treats the trip as the basic unit of observation, while the other considers a volume count as the basic unit of observation. Both approaches will be discussed in greater detail later, but for now it suffices to indicate that for any matrix with cells T_{ij}, the likelihood of the matrix can be estimated using a function $L = f(T_{ij}, t_{ij})$, where t_{ij} represents prior information. This prior information is often referred to as the seed matrix, and can be derived from either a previous study or survey. In the absence of such prior information, all of the cells in this prior matrix should be set to a uniform set of values.

In the case of the trip distribution process, the additional information that is added is usually, in the form of some expression of impedance. For example, the original gravity model considered that the likelihood of trips between two zones was proportional to the inverse of the square of the distance between the two zones. Since that time, many more sophisticated forms of impedance have been considered, but for the purposes of this discussion, all of these variations can be generalized as being of the form f_{ij}, where $f_{ij} = f(c_{ij})$ or the generalized cost of inter-zonal travel. What is less obvious, however, is the fact that the use of this set of impedance factors f_{ij}, is essentially equivalent to the use of a seed matrix t_{ij} as will be discussed later.

This implies that solving the trip distribution problem, using only zonal trip productions and attractions as constraints, together with a trip impedance matrix, is essentially the same as solving the synthetic O-D generation problem using zone connector in and out flows as constraints, and utilizing a matrix $t_{ij} = T*f_{ij}/(\Sigma f_{ij})$ as a seed matrix. This similarity is very encouraging to most traffic engineers and planners for the following reasons. First, many planners view the trip distribution process as a well-established approach to estimating O-D matrices, which has a long history and is hard to bring into question. In contrast, synthetic O-D generation is much less well known, and while many are not familiar with the details of the technique, most have heard of the difficulties arising from the under-specification problem. The above similarity indicates that planners should not question the validity of O-D matrices estimated synthetically any more than they now question the validity of O-D matrices estimated from the trip distribution process. Second, if planners are presently utilizing O-D matrices derived from a trip distribution exercise, they can now recognize that providing additional information about en-route links within a synthetic O-D generation process, can only make their matrix be more accurate, especially if they use the trip distribution matrix as a seed solution.

3.4 Maximum Likelihood for O-D Matrices Based on Trips

Figure 3.2 illustrates a simple network in which there are 2 origin zones connected to 2 destination zones. Some of the directional links connecting them carry only two O-D pairs each (links 1, 2, 4 and 5), while link 3 carries all four O-D pairs. If one considers the observed link flows in Table 3.3, one can note that a number of different matrices can be applied to this network and replicate the observed flows equally well. Table 3.4 shows that, in general, if the number of trips from A to C is considered to be x, that the number of trips between all other O-D pairs is automatically specified. The maximum likelihood approach, to selecting a synthetic O-D, therefore revolves around finding from all of these potential feasible options that matrix which is considered as the most likely.

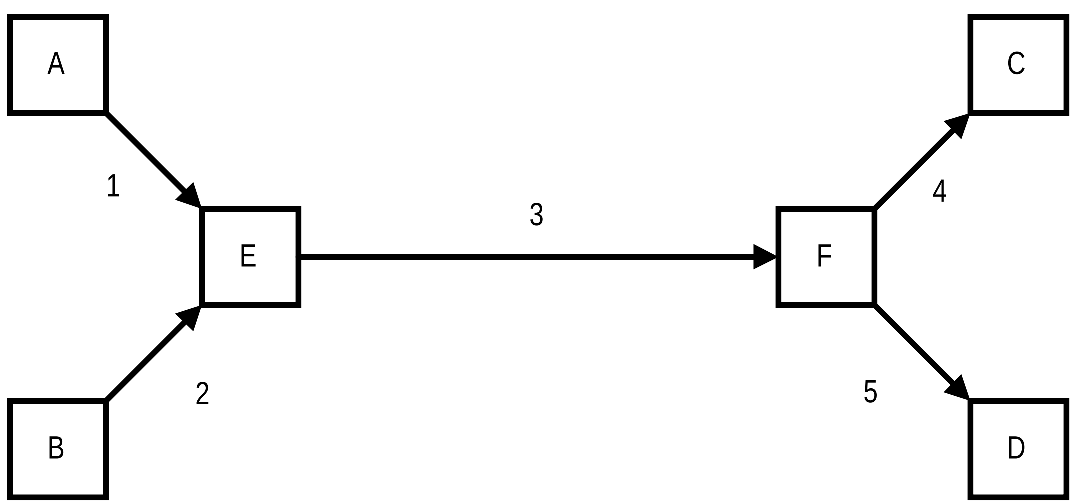

Figure 3.2: 5-Link Network

Table 3.3: Observed Link Flows

Link No.	Observed Volume
1	40
2	60
3	100
4	70
5	30

The trip-based approach, to defining maximum likelihood, considers that the overall trip matrix is made up of uniquely identifiable individual trip makers. In the case of Figure 3.2, and the matrix in Table 3.5, it can be noted that there are a total of 100 trips in the matrix.

Table 3.4: General Matrix

	C	D	Sum
A	X	40 – X	40
B	70 – X	X – 10	60
Sum	70	30	100

Table 3.5: Potential Matrix

	C	D	Sum
A	20	20	40
B	50	10	60
Sum	70	30	100

If one now considers the 20 trip makers from A to C, one can show that there are 100!/(20!*80!) unique ways of drawing 20 trip makers from the original population of 100. This leaves 80 trip makers to draw from for the 20 trip makers from A to D, implying that there are 80!/(20!*60!) different ways of doing this. Similarly, there are 60!/(10!*50!) ways of selecting trip makers from B to C, and 10!/(10!*0!) to select trip makers from B to D. In total, and after some simplification, this creates for 100!/(20!*20!*50!*10!) ways of creating a simple 20, 20, 50 and 10 trip matrix. In general, this number is: $T!/(\pi\ T_{ij}!)$. Figure 3.3 now shows, for all of the matrices identified in Table 3.4, their associated likelihood. It is clear from this figure that the most likely matrix is one where x = 28. The list of all possible solutions for the network in Figure 3.2 is given in Table 3.6

Maximize: $$Z(T_{ij}) = \frac{T!}{\prod_{ij}(T_{ij}!)} \tag{3.1}$$

The above formulation does not take into account the availability of any prior information, from for example a previous survey. If such a previous survey showed that the fractions of traffic, by O-D, are as indicated in Table 3.7, a seed matrix as shown in Table 3.8 could be constructed. While the seed matrix clearly does not satisfy the observed link flows, the seed matrix can be utilized to expand the maximum likelihood function as shown in Equation 3.2.

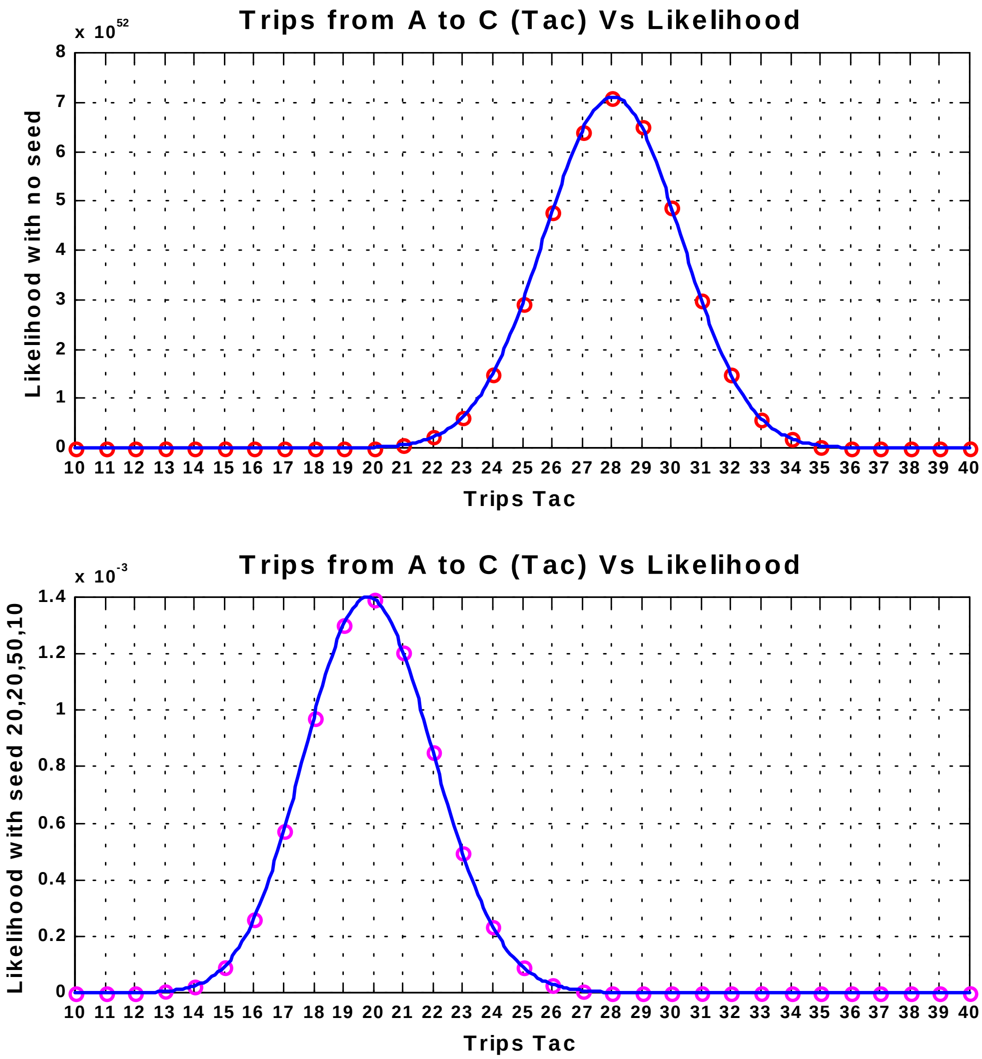

Figure 3.3: Trips from A to C vs. Likelihood With and Without a Seed Matrix

Table 3.6: Possible Solutions for the 5-Link Network

Demand Scenarios	T_{ac}	T_{ad}	T_{bc}	T_{bd}
0	10	30	60	0
1	11	29	59	1
2	12	28	58	2
3	13	27	57	3
4	14	26	56	4
5	15	25	55	5
6	16	24	54	6
7	17	23	53	7
8	18	22	52	8
9	19	21	51	9
10	20	20	50	10
11	21	19	49	11
12	22	18	48	12
13	23	17	47	13
14	24	16	46	14
15	25	15	45	15
16	26	14	44	16
17	27	13	43	17
18	28	12	42	18
19	29	11	41	19
20	30	10	40	20
21	31	9	39	21
22	32	8	38	22
23	33	7	37	23
24	34	6	36	24
25	35	5	35	25
26	36	4	34	26
27	37	3	33	27
28	38	2	32	28
29	39	1	31	29
30	40	0	30	30

Table 3.7: Matrix of Fraction of Traffic Obtained from Previous Survey

	C	D
A	0.1	0.2
B	0.3	0.4

Table 3.8: Seed Matrix based on fractions of traffic in Table 3.7

	C	D
A	10	20
B	30	40

It can be noted that the likelihood of an individual trip be from i to j is $t_{ij}/(\Sigma\ t_{ij})$,based on the above seed matrix. Consequently, the probability of T_{ij} trips being drawn is $(t_{ij}/(\Sigma\ t_{ij}))^{\wedge}T_{ij}$. It can be shown, in Figure 3.3, that the use of the above non-uniform seed matrix yields a different matrix as being identified as the most likely O-D matrix. Conversely, when a uniform seed matrix is applied to the network in Figure 3.2, it can be shown to have no impact. However, it will be shown later that this finding is unique to only certain network types, and does not hold in general. For reasons that will be discussed later, it will be considered that the formulation including the seed is the correct one, when the formulation with and without the seed differ.

Maximize: $$Z(T_{ij}, t_{ij}) = \frac{T!}{\prod_{ij}(T_{ij}!)} \prod_{ij} \left(\frac{t_{ij}}{\sum_{ij} t_{ij}} \right)^{T_{ij}} \qquad (3.2)$$

3.5 Maximum Likelihood Based on Volume Counts

An alternative maximum likelihood formulation considers that the basic unit of observation is a volume count. In the 5-link network of Figure 3.2, for example, it can be noted that there are 300 volume counts (30+70+100+40+60). If one now considers a potential solution matrix, such as the one in Table 3.7, one can show that the number of ways in which these volume counts can be drawn from the original set of 300 is as shown in Equation 3.3. If 40 is the link volume, and T_{ac} and T_{ad} are the trips which pass through the link, then the number of groups which can be obtained if 30 volume counts are divided between T_{ac} and T_{ad} is $40!/(T_{ac}!*T_{ad}!)$. Consequently, the probability of T_{ij} trips being drawn is $(t_{ij}/(\Sigma\ t_{ij}))^{\wedge}T_{ij}$. If this likelihood is computed for all the links then their product gives the net likelihood as shown in Equation 3.3. This formulation can also be derived based on Information Theory which is discussed in Chapter 4.

Maximize: $$Z\left(T_{ij}, t_{ij}\right) = \prod_a \frac{V_a!}{\prod_{ij}\left(T_{ij}\, p_{ij}^a\right)!} \prod_{ij} \left(\frac{t_{ij}\, p_{ij}^a}{v_a}\right)^{T_{ij}\, p_{ij}^a} \qquad (3.3)$$

Despite the somewhat different structure of Equation 3.3, it can be shown that for the network in Figure 3.2, this objective function identifies the same O-D matrix as being the most likely as either one of the two earlier trip based formulations. However, this is not a general finding. Specifically, in some cases where the trip based formulation without an explicit reference to a seed differs from the one which includes the seed, albeit a uniform one, the volume based formulation is only consistent with the latter solution (Equation 3.2). For this reason, the trip-based formulation, including the explicit accounting of the seed matrix, is viewed to be the bench mark against which all other solutions and formulations should be compared. A more detailed discussion of the various formulations is given in Chapter 4.

3.6 Presence of Flow Continuity

The above formulations of objective functions for expressing likelihood require additional constraints in order to be complete. The simplest of these constraints indicate that the sum of all the trips crossing a given link must be equal to the link flow on that link, as indicated in Equation 3.4. As will be shown later, the simplest mechanism, for including the above constraints in the earlier objective functions, is to utilize Lagrange multipliers. These multipliers permit an objective function with equality constraints to be transformed into an equivalent unconstrained objective function.

$$V_a = \sum_{ij} T_{ij}\, p_{ij}^a \qquad \forall\, a \qquad (3.4)$$

This simple set of equality constraints, while making the formulation complete, may at times also render the problem infeasible. For example, consider in the network of Figure 3.2 that the observed flows are 30, 70, 110, 40 and 60. In this case, there is no set of O-D matrices that can match the observed link flows, as there is no continuity of flow at the nodes that are interior to the network for these flows. Therefore, there would be no feasible solution space for either the objective function with constraints or the equivalent unconstrained optimization.

A more general formulation, therefore, is to request that the errors in the link flows are simply minimized, rather than being completely eliminated. In other words, rather than finding the most likely O-D matrix of all of those that exactly replicate the observed link flows, the problem is re-formulated as finding the most likely O-D matrix from among all of those that come equally close to matching the link flows.

One expression that captures the errors to be minimized is shown in Equation 3.5, and is subject to the flow continuity constraints in Equation 3.6. The constraints in Equation 3.6 can be introduced in Equation 3.5 to yield an unconstrained objective function, yielding a set of complementary link flows V_a'. These complementary flows are those which deviate the least from the observed link flows, while still satisfying link flow continuity. Given that these complementary link flows do satisfy link continuity, they can now be added as rigid equality constraints to the objective function in either Equation 3.1 or 3.2, and can be guaranteed to yield a feasible solution.

Minimize: $$Z(T_{ij}) = \sum_a \left(V_a - V_a'\right)^2 \qquad \forall\, a \tag{3.5}$$

Where:

V_a Actual observed link volume

V_a' Volumes that are closest to V_a and satisfy flow continuity

$$V_a' = \sum_{ij} T_{ij}\, p_{ij}^a \qquad \forall\, a \tag{3.6}$$

Alternatively, one can take the Equations in 3.6 and substitute them into Equation 3.5. This new expression, which is shown as Equation 3.7, is one which should be minimized concurrently to maximizing the objective function in Equation 3.1 or 3.2.

Unfortunately, it is not easy to combine one expression, that desires to maximize likelihood, with another that desires to minimize link flow error, as Lagrangian can only add equality constraints to a constrained objective function.

A solution to this problem involves taking the partial derivatives of Equation 3.7 with respect to each of the trip cells that are to be estimated. This yields, as shown in Equation 3.8, as many equations as there are trip cells. Furthermore, setting these derivatives equal to 0 is equivalent to minimizing Equation 3.7. However, while Equation 3.7 could not be added to the maximum likelihood objective function, the equalities in Equation 3.8 can. This produces an unconstrained objective function that will always yield a feasible solution, as is shown in Equation 3.9.

$$\text{Minimize: } Z(T_{ij}) = \sum_a \left(V_a - \sum_{ij} T_{ij} p_{ij}^a \right)^2 \qquad \forall\, a \tag{3.7}$$

$$0 = 2\left(\sum_a \left(V_a \cdot p_{ij}^a\right) - \left(\sum_a p_{ij}^a \left(\sum_{ij} T_{ij} p_{ij}^a \right)\right)\right) \qquad \forall\, i, j \tag{3.8}$$

$$L \equiv \frac{T!}{\prod_{ij} T_{ij}} \prod_{ij} \left(\frac{t_{ij}}{t} \right)^{T_{ij}} - \sum_{ij} \left(\lambda_{ij} \cdot 2\left(\sum_a \left(V_a \cdot p_{ij}^a\right) - \left(p_{ij}^a \sum_a \left(\sum_{ij} T_{ij} p_{ij}^a \right)\right)\right)\right) \qquad \forall\, i, j \tag{3.9}$$

The net result, of the above process, is to suggest that most synthetic O-D generation problems consist of 2 sub problems. One of these involves finding a new set of complementary link flows that do permit link flow continuity to exist, at which point the maximum likelihood problem can be solved as before. Alternatively, one can compute the partial derivatives, that will yield link flow continuity, while deviating by the least amount from the observed link flows. These partial derivatives can then be utilized directly in the maximum likelihood formulation using Lagrangian multipliers. Both solutions can be shown to yield identical results.

A first challenge, to maximizing Equation 3.9, is that it yields numbers of a very large magnitude that are often very difficult to work with, while maintaining adequate precision. Furthermore, as it is common to maximize objective functions by taking their derivatives, and as it is more difficult to contemplate the derivative of a discontinuous expression, (such as those including factorials) a simple approximation is made. This approximation involves taking the natural logarithm of either objective function Equation 3.1 or 3.2. Taking the log of the objective function both makes the output easier to handle, and allows us to use Stirling's approximation as a convenient continuous equivalent to the $Ln(x!)$.

The resulting converted objective functions (based on original objective function 3.2) is illustrated as Equation 3.10. In addition, Equation 3.11 illustrates the derivatives of this log-transformed objective function. When these derivatives of log-transformed objective functions are augmented with the previously mentioned partial derivatives that minimize link flow error, Equation 3.12 emerges. This equation, when solved, yields a the most likely O-D matrix of all of those matrices that come equally close to matching the observed link flows.

The above set of non-linear equations can be solved numerically in a number of ways using a variety of standard numerical analysis software packages. However, as for large networks the number of equations and unknowns becomes prohibitive, a special purpose equation solver has been developed, called QUEENSOD. This solver fully optimizes the objective function in either Equation 3.2 or 3.3, subject to also minimizing the link flows when link flow continuity does not exist. The only approximation that it does make is Stirling's, which has been shown to produce errors less than 1% for the range of values and derivatives being typically considered in synthetic O-D generation. A detailed derivation of Equation 3.12 is given in Chapter 5.

$$\text{Maximize: } \begin{array}{l} T\, ln\left(\frac{T}{t}\right) - T - \sum_{ij}\left(T_{ij}\, ln\left(\frac{T_{ij}}{t_{ij}}\right) - T_{ij}\right) - \\ \sum_{ij}\left(\lambda_{ij} \cdot 2\left(\sum_{a}\left(V_a \cdot p_{ij}^a\right) - \left(p_{ij}^a \sum_{a}\left(\sum_{ij} T_{ij}\, p_{ij}^a\right)\right)\right)\right) \quad \forall\, i, j \end{array} \tag{3.10}$$

$$0 = Ln(T) - Ln(t) - Ln(T_{ij}) + Ln(t_{ij}) \qquad \forall\, i, j \tag{3.11}$$

$$\begin{array}{l} 0 = Ln(T) - Ln(t) - Ln(T_{ij}) + Ln(t_{ij}) - \\ \sum_{ij}\left(\lambda_{ij} \cdot 2\left(\sum_{a}\left(V_a \cdot p_{ij}^a\right) - \left(\sum_{a} p_{ij}^a\left(\sum_{ij} T_{ij}\, p_{ij}^a\right)\right)\right)\right) \quad \forall\, i, j \end{array} \tag{3.12}$$

3.7 Other Approximations

While QUEENSOD solves the full objective function, utilizing only Stirling's approximation, many others have made more drastic approximations. Some of these, such as those by Willumsen, are widely utilized. However, they can also be shown to introduce considerable error, as will be indicated next.

The most common approximation that is applied, after Stirling's approximation is invoked, is to consider that the total number of trips in the network T is constant, and then to drop this term from the optimization. This yields Equation 3.13. Unfortunately, as is shown in Figure 3.5 for the network illustrated in Figure 3.4 (whose link volumes are given in Table 3.9) dropping T for networks in which the total number of trips is not constant, can yield a very different selection of the "optimum" or "most likely" matrix.

$$\text{Maximize: } Z(T_{ij}, t_{ij}) = -\sum_{ij}\left(T_{ij} Ln\left(\frac{T_{ij}}{t_{ij}}\right) - T_{ij}\right) \tag{3.13}$$

A detailed derivation of the various formulations is provided in the following Chapter.

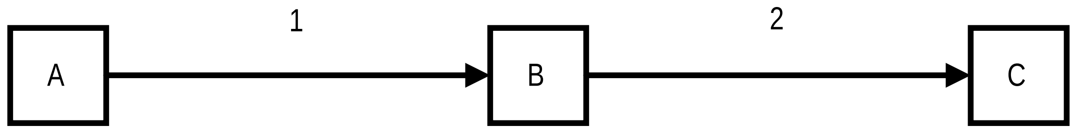

Figure 3.4: 2-Link Network

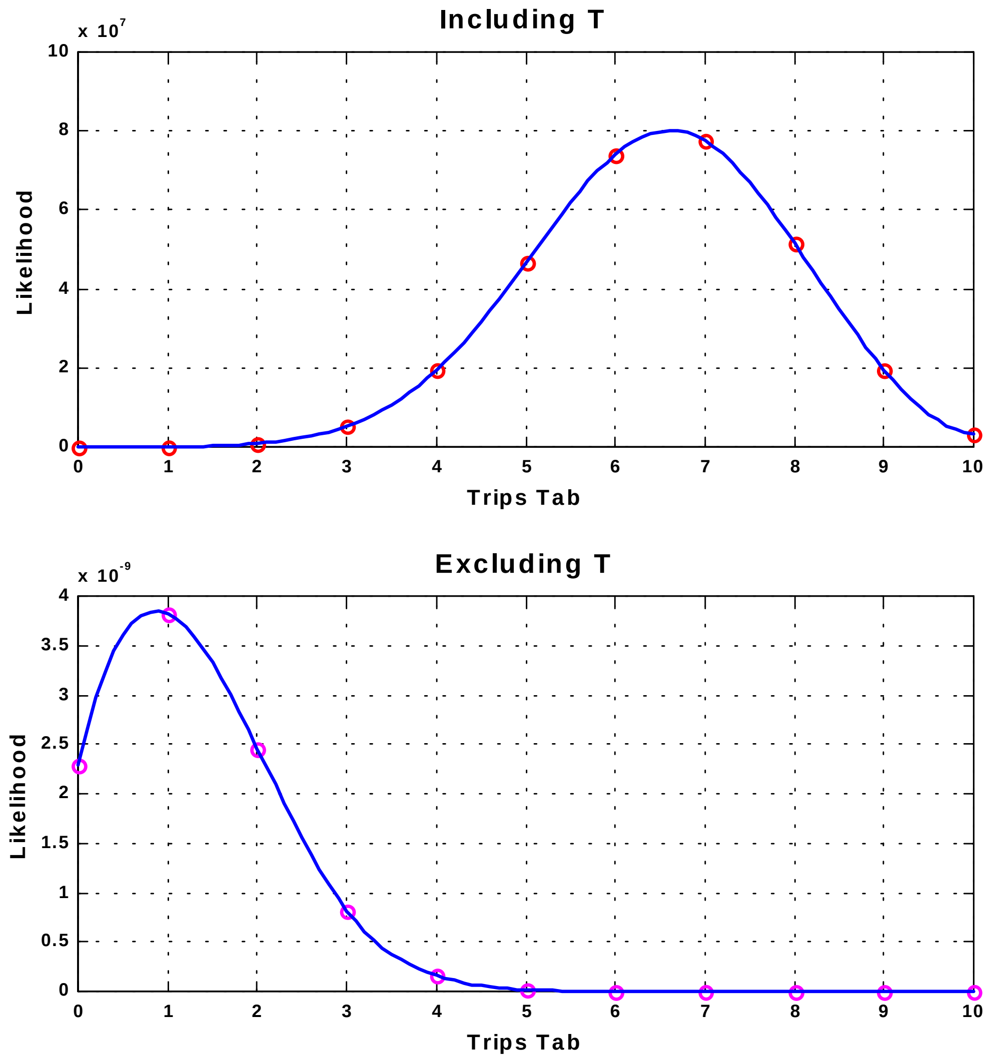

Figure 3.5: Comparison of Optimal Solution; Including T and Excluding T

Table 3.9: Link Volume on the 2 Link Network

Link Number	Volume
1	10
2	15

3.8 Summary

In this Chapter an overview of various sub problems was presented. Some of the sub problems which one must deal with is, the lack of flow continuity and presence of multiple solutions to the problem. The flow continuity problem is often solved by redistributing the errors to obtain an equivalent link volume which posses flow continuity. The problem of multiple solutions is solved, as explained, by using a priori matrix. This still does not completely answer the question of multiple solutions but at least narrow downs the solution search.

Similarities between trip distribution and synthetic O-D generation was presented and an overview of maximum likelihood approach based on trips and volume counts was also discussed. A detailed description and derivations of the various formulation which were mentioned in this Chapter is presented in the following Chapter.