

A Strange Property of TSP and its Consequences

Seenu S. Reddi
ReddiSS@AOL.COM

Abstract: A strange property of the traveling salesman problem hitherto unknown is presented and its consequences in finding the near optimal schedules are described.

This note describes a finding of an unknown property of the well-known Traveling Salesman Problem (TSP) and the implications of this property. The property can be roughly stated as that given a TSP with its distance matrix, another new distance matrix can be found from this matrix such that *any* tour has the *same* distance. We first give an example and then present the mathematical basis for derivation of this property.

Consider a 4-node TSP with the distance matrix:

$$\mathbf{D} = \begin{bmatrix} 0.000 & 51.000 & 477.000 & 344.000 \\ 510.000 & 0.000 & 179.000 & 734.000 \\ 488.000 & 368.000 & 0.000 & 972.000 \\ 474.000 & 715.000 & 155.000 & 0.000 \end{bmatrix}$$

We derive a new distance matrix:

$$\mathbf{M} = \begin{bmatrix} 0.000 & 202.959 & 132.459 & 536.585 \\ 425.210 & 0.000 & 296.834 & 700.960 \\ 536.710 & 478.835 & 0.000 & 812.460 \\ 510.085 & 452.209 & 381.709 & 0.000 \end{bmatrix}$$

Tours and their distances for D and M are computed:

Tour	M	D
1 2 3 4	1822.34	1676
1 2 4 3	1822.34	1428
1 3 2 4	1822.34	2053
1 3 4 2	1822.34	2674
1 4 2 3	1822.34	1726
1 4 3 2	1822.34	1377

As can be noted the tour distance varies from a minimum of 1377 (optimum) to a maximum of 2674 whereas the tours of M have a constant distance of 1822.34. To obtain the optimal tour a difference matrix can be formed $DM = D - M$:

$$\mathbf{DM} = \begin{bmatrix} 0.000 & -151.959 & 344.541 & -192.585 \\ 84.790 & 0.000 & -117.834 & 33.040 \\ -48.710 & -110.835 & 0.000 & 159.540 \\ -36.085 & 262.791 & -226.709 & 0.000 \end{bmatrix}$$

The optimal tour 1-4-3-2-1 has edges of the value -192.585 , -226.709 , -110.835 and 84.790 from which we get the optimal distance of 1377 from the tour constant 1822.34.

Before we proceed to mathematical basis, certain observations can be made. From the difference matrix, selecting negative elements decreases the tour cost and the negative elements abound as many as positive ones. Also the constant tour distance 1822 is approximately in between the maximum of 1377 and 2624.

The mechanics of deriving the approximate matrix M from D are as follows. It is well known that the TSP formulated by Gilmore and Gomory^{1,2} is solvable polynomially and the distance matrix is given by $d_{i,j} = \max\{b_j - a_i, 0\}$. A two-machine flowshop problem with no wait can be formulated that leads to the solvable TSP². Our motivation is to extract a solvable case from the given TSP and define variables $\{a_i, b_i\}$ to approximate $\{d_{i,j}\}$ by setting up an equation system: $b_j - a_i = d_{i,j}$. Thus for a n -node TSP, we define n -job, 2-machine problem with processing times $\{a_i, b_i\}$, $i = 1, 2, \dots, n$ approximate the transition time between job i and job j to be $d_{i,j}$. For the sake of exposition, we assume a 4-node TSP. We define $\mathbf{d} = [d_{1,2}, d_{1,3}, d_{1,4}, \dots, d_{4,1}, d_{4,2}, d_{4,3}]$ (we indicate column vectors by $[..]$ and row vectors $(..)$) and $\mathbf{f} = [a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4]$. Note \mathbf{d} has 12 components and \mathbf{f} has 8. Matrix \mathbf{R} with dimensions of 12×8 is set up to transform \mathbf{f} to \mathbf{d} :

$$\mathbf{Rf} = \mathbf{d} \quad (1)$$

The first row of \mathbf{R} will be $(-1, 0, 0, 0, 0, 1, 0, 0)$ since $d_{1,2} = b_2 - a_1$ and so on. Let \mathbf{S} be the pseudoinverse of \mathbf{R}^3 and (1) can be rewritten as:

$$\mathbf{f} = \mathbf{Sd} \quad (2)$$

where

$$\mathbf{S} = \begin{bmatrix} -0.292 & -0.292 & -0.292 & -0.042 & 0.083 & 0.083 & -0.042 & 0.083 & 0.083 & -0.042 & 0.083 & 0.083 \\ -0.042 & 0.083 & 0.083 & -0.292 & -0.292 & -0.292 & 0.083 & -0.042 & 0.083 & 0.083 & -0.042 & 0.083 \\ 0.083 & -0.042 & 0.083 & 0.083 & -0.042 & 0.083 & -0.292 & -0.292 & -0.292 & 0.083 & 0.083 & -0.042 \\ 0.083 & 0.083 & -0.042 & 0.083 & 0.083 & -0.042 & 0.083 & 0.083 & -0.042 & -0.292 & -0.292 & -0.292 \\ 0.042 & 0.042 & 0.042 & 0.292 & -0.083 & -0.083 & 0.292 & -0.083 & -0.083 & 0.292 & -0.083 & -0.083 \\ 0.292 & -0.083 & -0.083 & 0.042 & 0.042 & 0.042 & -0.083 & 0.292 & -0.083 & -0.083 & 0.292 & -0.083 \\ -0.083 & 0.292 & -0.083 & -0.083 & 0.292 & -0.083 & 0.042 & 0.042 & 0.042 & -0.083 & -0.083 & 0.292 \\ -0.083 & -0.083 & 0.292 & -0.083 & -0.083 & 0.292 & -0.083 & -0.083 & 0.292 & 0.042 & 0.042 & 0.042 \end{bmatrix}$$

Once \mathbf{f} is computed from (2), the new distance matrix \mathbf{M} can be computed:

$$\mathbf{M} = [m_{i,j}] = [b_j - a_i] \quad (3)$$

It is simple to prove that the tours of \mathbf{M} are of the same length since all the tour lengths will be given by $\sum b_i - \sum a_i$ where the summation is over all the n jobs.

Since (2) represents a least squares solution, it can be expected that approximately half the elements in \mathbf{DM} will be positive and the other half negative. The numerical simulation conducted on the 4-city case bears out this observation.

The above procedure can be generalized to n -node TSP and \mathbf{S} in this case will be of dimensions $n(n-1) \times 2n$. We conjecture that the constant tour length of \mathbf{DM} will be half way between the extrema of \mathbf{D} in a statistical sense and that \mathbf{DM} will have approximately half of its elements negative. Near-optimal tours can be found by using greedy-type algorithms on this set of negative elements with polynomial complexity. This is an internet publication and can be referenced as:

Reddi, S. S., "A Strange Property of TSP and its Consequences," www.rspq.org/pubs, 2004.

References:

- ¹ P. C. GILMORE AND R. E. GOMORY (1964) Sequencing a one state-variable machine: a solvable case of traveling salesman problem. *Ops. Res.* **12**, 655.
- ² S. S. REDDI AND C. V. RAMAMOORTHY (1972) On the flow-shop sequencing problem with no wait in process. *Opl Res Q.*, **23**, 323.
- ³ G. H. GOLUB AND C. F. VAN LOAN (1983) *Matrix Computations*. Johns Hopkins University Press, Baltimore, Md.