

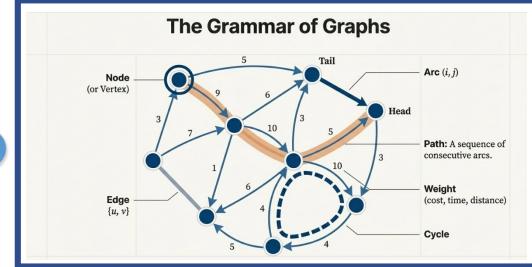
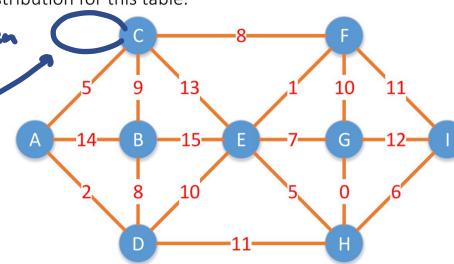
C1 Introduction to Transport

Question 1

Consider the weighted graph $G = (V, E)$ illustrated in the figure below.

- Calculate the degree of every node. How about indegrees and outdegrees?
- Provide the node adjacency matrix for the graph, and a forward star representation that would include link weights.
- Can an adjacency matrix be used to calculate the degrees of a node i ? Provide a formula.
- How would you represent a buckle added to node C in the adjacency matrix?
- Obtain a degree distribution for this table.

info: (d) a buckle is when a node is connected to itself with one edge.



Solution 1a

This is an undirected graph – as such, nodes only have degrees.

Node	Degree
A	3
B	4
C	4
D	4
E	6
F	4
G	4
H	4
I	3

degree is no. of edges connected

indegree and outdegrees are for directed graph.

Solution 1b

The node adjacency matrix for the graph is as follows:

	to								
from	A	B	C	D	E	F	G	H	I
A	0	1	1	1	0	0	0	0	0
B	1	0	1	1	1	0	0	0	0
C	1	1	0	0	1	1	0	0	0
D	1	1	0	0	1	0	0	1	0
E	0	1	1	1	0	1	1	1	0
F	0	0	1	0	1	0	1	0	1
G	0	0	0	0	1	1	0	1	1
H	0	0	0	1	1	0	1	0	1
I	0	0	0	0	0	1	1	1	0

two different ways to represent graph with tables

show weight

A forward star representation for the graph is as follows:

Node	(continued)		
	Node1	Node2	Cost
A	A	C	5
B	C	A	5
C	A	B	14
D	B	A	14
E	A	D	2
F	D	A	2
G	C	B	9
H	B	C	9
I	C	E	13
	E	C	13
	C	F	8
	F	C	8
	B	E	15
	E	B	15
	D	B	8

(continued)		
Node1	Node2	Cost
B	D	8
D	E	10
E	D	10
D	H	11
H	D	11
E	F	1
F	E	1
E	H	5
H	E	5
E	G	7
G	E	7
F	G	10
G	F	10
G	H	0
H	G	0

- if this is directed graph there will not be $A \rightarrow C$ and $C \rightarrow A$ so overall the table would be half the size.

Solution 1c

For every node $i \in V$, the degree will be calculated from the following function, where $k_{i,j}$ is the value of the adjacency matrix in position i, j :

$$\sum_{j \in V} k_{i,j}$$

Solution 1d

Since this graph is undirected, the adjacency matrix is transposable – a value of 1 on the cell C,C would not capture the bidirectional nature of the buckle. Instead, the convention is to use a value of 2.

this is not weight! this just show undirected graph C→C should count as two!

Solution 1e

The degree distribution for the graph is as follows:

Degree	Occurrence	Probability
1	0	0
2	0	0
3	2	0.222
4	6	0.667
5	0	0
6	1	0.111

Question 2

Suppose that every node in a directed graph has a nonzero indegree. Show that the graph must contain a directed cycle. Hint: use an example graph in your response.

Solution 2

Let i_1 be any arbitrary node in the graph. Since i_1 has a positive indegree, there exists some node i_2 such that $(i_2, i_1) \in A$, where A is the set of all edges.

Again, since i_2 has a positive indegree, there exists some node $(i_3, i_2) \in A$. If $i_3 = i_1$, a directed cycle has just been discovered in the graph; otherwise we continue this process until we visit a node twice.

Notice that i_1 will eventually be reached since the graph has a finite number of nodes. In this case, we have discovered a directed cycle in the graph.

C2. Network Algorithm

Network Algorithm.

Shortest Path Problem

Dijkstra's Algorithm

Clustering Problem.

- to identify group with similar characteristic within a collection of element.

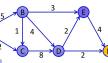
A^* Algorithm.

- Modification of Dijkstra's (used for larger network)
- Differences: Instead of checking all nodes, it select next node based on least estimated total cost. (a heuristic approach)

Tabular Implementation Step 1		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	∞ A	
C	∞ A	
D	∞ A	
E	∞ A	
F	∞ A	



Tabular Implementation Step 5		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	∞ A A	
C	∞ A A	
D	∞ A A C	
E	∞ A A A B	
F	∞ A A A A E	$B+4=12$



Tabular Implementation Step 2



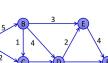
Tabular Implementation Step 6		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	∞ 5 A	5 A
C	∞ 2 A	2 A
D	∞ A A	
E	∞ A A	
F	∞ A A	



Tabular Implementation Step 3



Tabular Implementation Step 7		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	0 A
B	∞ 5 A	5 A
C	∞ 2 A	2 A
D	∞ A A	C
E	∞ A A	B
F	∞ A A	A



Tabular Implementation Step 4



Tabular Implementation Step 8	
Node	Final
A	0 A
B	5 A
C	2 A
D	9 B
E	8 B
F	11 D



Final path cost (label): 11
Route: F \leftarrow D \leftarrow B \leftarrow A



Tabular Implementation Step 9		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	∞ 5 A	5 A
C	∞ 2 A	2 A
D	∞ 9 B	B
E	∞ 8 B	
F	∞ 11 D	



Tabular Implementation Step 10		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 11		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 12		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 13		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



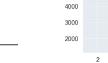
Tabular Implementation Step 14		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 15		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 16		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 17		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 18		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	



Tabular Implementation Step 19		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	

Tabular Implementation Step 20		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	

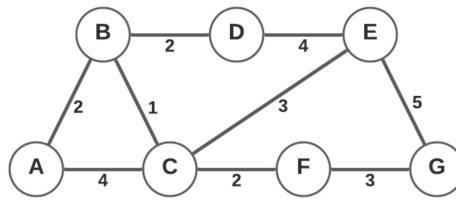
Tabular Implementation Step 21		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	

Tabular Implementation Step 22		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C	2 A	
D	9 B	
E	8 B	
F	11 D	

Tabular Implementation Step 23		
Node	Loose end table (tentative labels and backnodes)	Final
A	0 A	
B	5 A	
C		

Question 2 (how to use Dijkstra's to find closeness centrality)

The figure below illustrates a network graph. b. Use Dijkstra's algorithm to calculate the closeness centrality for vertex D.



Solution 2

start from D →

Node	Loose End Table A to G							Final
	∞	∞	4	4	D → B → A = 4	4	B	
A	∞	∞	4	4	D → B → A = 4	4	B	4
B	∞	2	D	—	D → B → C → A = 7	7	D	2
C	∞	∞	3	D	—	3	D	3
D	0	—	D	—	—	0	D	0
E	∞	4	D	D	D → E = 4	4	D	4
F	∞	∞	∞	5	C	5	C	5
G	∞	∞	∞	∞	9	E	8	8

Node	Loose End Table A to G							Final
	∞	∞	4	4	4	4	4	
A	∞	∞	4	4	B	—	—	4
B	∞	2	D	—	—	—	—	2
C	∞	∞	3	D	—	—	—	3
D	0	—	D	—	—	—	—	0
E	∞	4	D	D	D	D	—	4
F	∞	∞	∞	5	C	C	—	5
G	∞	∞	∞	∞	9	E	8	8

but since A only A → B, A → C
and B and C already expanded,
A is ended.

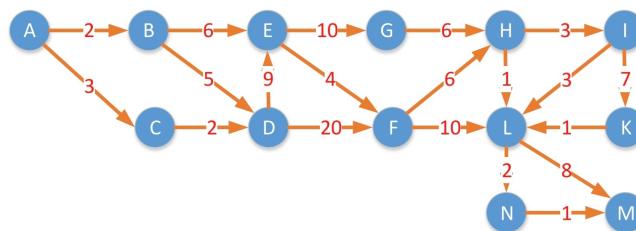
Node	Loose End Table A to G							Final
	∞	∞	4	4	4	4	4	
A	∞	∞	4	B	—	—	—	4
B	∞	2	D	—	—	—	—	2
C	∞	∞	3	D	—	—	—	3
D	0	—	D	—	—	—	—	0
E	∞	4	D	D	D	D	—	4
F	∞	∞	∞	5	C	C	—	5
G	∞	∞	∞	∞	9	E	8	8

The average distance from D to the other nodes in the network is 4.333 and therefore the closeness centrality of node D is 0.23.

$$\text{closeness centrality} = \frac{1}{\text{average distance}} = \frac{1}{4.333} = 0.23$$

Question 3

- a. Using Dijkstra's algorithm, determine the shortest path between nodes A and M for the graph below.
- b. Describe the steps involved in the algorithm that would find the shortest path between an origin and destination pair, subject to the additional condition that the path must visit a specified node.
- c. For the graph shown below, find a path between A and M that also visits node G.



Solution 3a

We set up a loose end table with all the nodes and traverse the network using the Dijkstra's algorithm.

Node	Loose End Table												Final
A	0												0
A	.												A
B	∞	$0+2=2$											2
A	A	.											A
C	∞	$0+3=3$	3										3
A	A	A	.										A
D	∞	∞	$2+5=7$	$3+2=5$									5
A	A	B	C	.									C
E	∞	∞	$2+6=8$	8	8	.							8
A	A	B	B	B	.								B
F	∞	∞	∞	∞	$5+20=25$	$8+4=12$.						12
A	A	A	A	D	E	.							E
G	∞	∞	∞	∞	$8+10=18$	18	.						18
A	A	A	A	E	E	.							E
H	∞	∞	∞	∞	∞	$12+6=18$	18	.					18
A	A	A	A	A	F	F	.						F
I	∞	∞	∞	∞	∞	∞	$18+3=21$	21	.				21
A	A	A	A	A	A	A	H	H	.				H
K	∞	∞	∞	∞	∞	∞	∞	∞	$21+7=28$	28	28	.	28
A	A	A	A	A	A	A	A	I	I	I	.	I	
L	∞	∞	∞	∞	∞	$12+10=22$	22	$18+1=19$.				19
A	A	A	A	A	A	F	H	.					H
M	∞	∞	∞	∞	∞	∞	∞	$19+8=27$	27	$21+1=22$.		22
A	A	A	A	A	A	A	A	L	L	N	.	N	
N	∞	∞	∞	∞	∞	∞	∞	$19+2=21$	21	.			21
A	A	A	A	A	A	A	A	L	L	.			L

According to the final column, the total cost of travelling from A to M is 22. If we backtrack from M using the predecessor labels in the final column, the optimal path is: M \leftarrow N \leftarrow L \leftarrow H \leftarrow F \leftarrow E \leftarrow B \leftarrow A

Solution 3b

Many approaches could be used to solve this problem – an acceptable solution would suggest that we specify an algorithm that finds the shortest path from A to G and another from G to M. The final path would be a combination of the two paths.

Solution 3c

Node	Loose End Table A to G						Final
A	0						0
A	.						A
B	∞	$0+2=2$	A				2
A	.						A
C	∞	$0+3=3$	3	A			3
A	.						A
D	∞	∞	$2+5=7$	$3+2=5$	C		5
A	A	B	B	C	.		C
E	∞	∞	$2+6=8$	8	8	.	8
A	A	B	B	B	.		B
F	∞	∞	∞	∞	$5+20=25$	$8+4=12$	12
A	A	A	A	D	E	.	E
G	∞	∞	∞	∞	∞	$8+10=18$	18
A	A	A	A	A	E	.	E

Node	Loose End Table G to M						Final
G	0						0
G	.						G
H	∞	$0+6=6$	H				6
G	.						H
I	∞	∞	$6+3=9$	9	9	.	9
G	H	H	H	H	.		H
K	∞	∞	$6+1=7$	G	G	$9+7=16$	16
G	G	H	G	G	G	I	I
L	∞	∞	$6+1=7$	H			7
G	G	H	G	.			H
M	∞	∞	$7+8=15$	L	$9+1=10$	10	10
G	G	G	L	N	N	N	N
N	∞	∞	$7+2=9$	L			9
G	G	G	L	.			L

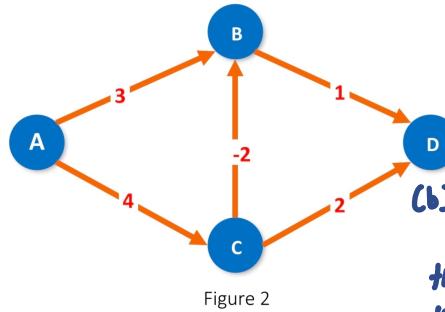
From A to G: A \rightarrow B \rightarrow E \rightarrow G, cost: 18

From G to M: G \rightarrow H \rightarrow L \rightarrow N \rightarrow M, cost: 10

Path between A and M, that also visits node G: A \rightarrow B \rightarrow E \rightarrow G \rightarrow H \rightarrow L \rightarrow N \rightarrow M, total cost: 28

Question 4

- Apply Dijkstra's algorithm to obtain the shortest path between nodes A and D at the weighted digraph provided in Figure 2.
- Comment on your findings and suggest an additional step in your approach that would address any issues that you have encountered.
- Undertake your own research to identify an algorithm that is capable of determining paths in graphs with negative edge weights.



(b) suggested step:
if found a shorter path
that already visited, mark as
unvisited again and add back to
processing queue

(a)

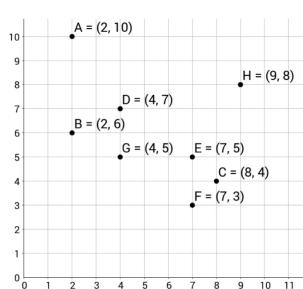
A	0	-	
B	0	3	-
C	0	4	-
D	0	0	-

expend C as it ties with D
 $C \rightarrow D = 6 < 4$
 (does not replace)

can't revisit B as already close case.

Question 6 (K-Mean Question)

A contractor has 8 project sites with the coordinates and Euclidian distances given in the table below. The company has decided that 2 distribution centres will be established to supply these sites with materials.



Site	X	Y	d(A)	d(B)	d(C)	d(D)	d(E)	d(F)	d(G)	d(H)
A	2	10	0.0	4.0	8.4	3.6	7.0	8.6	5.3	7.2
B	2	6	0.0	6.3	2.2	5.1	5.8	2.2	7.2	8
C	8	4		0.0	5.0	1.4	1.4	4.1	4.1	2
D	4	7		0.0	3.6	5.0	2.0	5.1		0
E	7	5			0.0	2.0	3.0	3.6		1
F	7	3				0.0	3.6	5.3		9
G	4	5					0.0	5.8		3
H	9	8						0.0		0

- Use a pen-and-paper (and calculator) approach to follow the steps involved in the k-means clustering algorithm to obtain two clusters for the above points. You can start with points A and C as initial guess for the centre points.

Solution 6a

Step 1:

Site	d(P1)	d(P2)
A	0.00	8.49
B	4.00	6.32
C	8.49	0.00
D	3.61	5.00
E	7.07	1.41
F	8.60	1.41
G	5.39	4.12
H	7.28	4.12

Cluster 1: {A, B, D}
 Cluster 2: {C, E, F, G, H}

New centre points:

$$P1 = \left(\frac{2+2+4}{3}, \frac{10+6+7}{3} \right) = (2.67, 7.67)$$

$$P2 = \left(\frac{8+7+7+4+9}{5}, \frac{4+5+3+5+8}{5} \right) = (7.00, 5.00)$$

Step 2:

Site	d(P1)	d(P2)
A	2.43	7.07
B	1.80	5.10
C	6.47	1.41
D	1.49	3.61
E	5.09	0.00
F	6.37	2.00
G	2.98	3.00
H	6.34	3.61

Cluster 1: {A, B, D, G}
 Cluster 2: {C, E, F, H}

New centre points:

$$P1 = \left(\frac{2+2+4+4}{4}, \frac{10+6+7+5}{4} \right) = (3.00, 7.00)$$

$$P2 = \left(\frac{8+7+7+9}{4}, \frac{4+5+3+8}{4} \right) = (7.75, 5.00)$$

Step 3:

Site	d(P1)	d(P2)
A	3.16	7.62
B	1.41	5.84
C	5.83	1.03
D	1.00	4.25
E	4.47	0.75
F	5.66	2.14
G	2.24	3.75
H	6.08	3.25

The cluster partition has not changed with respect to the previous step. Therefore, the optimal clusters are:

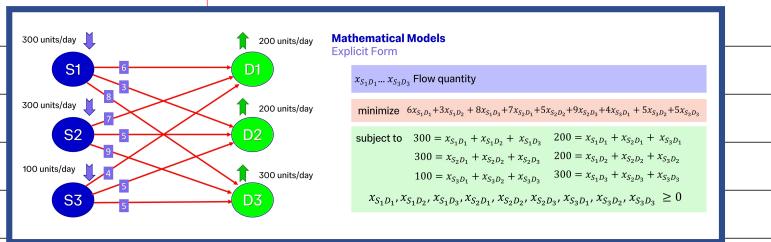
Cluster 1: {A, B, D, G}
 Cluster 2: {C, E, F, H}

c3 Mathematical Models (Minimise Cost / Maximise Profit)

Mathematical Models

What is it?

1. Minimum cost problem / Maximise Profit (Supply Demand)



How to solve?

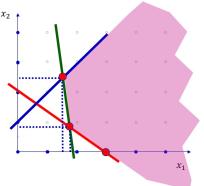
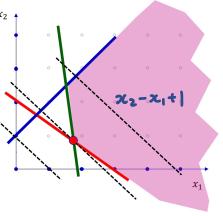
Method 1:

Constructing an LP graph

$$\begin{aligned} & \min x_1 + x_2 \\ & 2x_1 + 3x_2 \geq 6 \\ & -x_1 + 2x_2 \geq 2 \\ & 5x_1 + x_2 \geq 10 \end{aligned}$$

$$\begin{aligned} & \min z_1 + z_2 \\ & \text{plot } z_1 + z_2 = k \quad (\text{let } k=1) \\ & z_2 = -z_1 + 1 \end{aligned}$$

Method 2: Constructing an LP graph



(min cost pro. without transit)

S	Set of nodes
E	Set of transportation links
x_{ij}	Flow of goods between i, j
C_{ij}	Cost of transport on arc (i, j)
D_i	Cargo supply/demand on vertex i
minimize	$\sum_{(i,j) \in E} C_{ij} x_{ij}$
subject to	$\sum_{j \in N} x_{ij} - \sum_{j \in N} x_{ji} = D_i \quad \forall i \in N$
	$0 \leq x_{ij} \leq U_{ij} \quad \forall i, j \in E$

Transhipment Symbolic formulation

(min cost pro. with transit)

- The modified representation of the network allows us to use a significantly simpler model formulation.
- A single constraint now captures production, consumption and transhipment

N	Set of nodes
E	Set of transportation links
x_{ij}	Flow of goods between i, j
C_{ij}	Cost of transport on arc (i, j)
D_i	Cargo supply/demand on vertex i
minimize	$\sum_{(i,j) \in E} C_{ij} x_{ij}$
subject to	$\sum_{j \in N} x_{ij} - \sum_{j \in N} x_{ji} = D_i \quad \forall i \in N$
	$0 \leq x_{ij} \leq U_{ij} \quad \forall i, j \in E$

out - in = capacity

capacity of supply side
capacity of demand-side
transhipment node usually 0

can be:

- linear (continuous)
- integer (discrete)
- boolean (0 or 1)

can be:

- single: $\max f(x)$ or $\min f(x)$
- multiple: $\max f(x)$ and $\min g(x)$

(can be written as:
 $\max\{Af(x) - Bg(x)\}$)

Maximin or Minimax: $\min(\max f(x))$

(can be written as:
 $\min(f(x) - z)$)

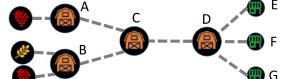
in this example we only transhipping one type of good.

3. For multicommodity flow

- An variable L_{ij} has been added to represent the capacity limit on each arc.
- A new constraint imposes an upper limit on the sum of flows within a link.

N	Set of vertices
E	Set of arcs
P	Set of commodities
x_{ij}^c	Flow of commodity c on arc (i, j)
C_{ij}	Cost of transport on arc (i, j)
D_i^c	Production of commodity c on vertex i
L_{ij}	Capacity limit on arc (i, j)
minimize	$\sum_{c \in P} \sum_{(i,j) \in E} C_{ij} x_{ij}^c$
subject to	$\sum_{j \in N} x_{ij}^c - \sum_{j \in N} x_{ji}^c = D_i^c \quad \forall i \in N, c \in P$
	$\sum_{c \in P} x_{ij}^c \leq L_{ij} \quad \forall i, j \in E$
	$0 \leq x_{ij}^c \leq U_{ij} \quad \forall i, j \in E, c \in P$

we minimise total cost (this inc. summing all c)
we track flow cons. separate (eg. grapes in grape out)



but for capacity limit, each link limit doesn't care what commodity it carries but so sum of all the c must adhere the capacity limit.

(eg. a train doesn't care it carries apple or orange but it can't exceed its capacity)

2. Knapsack Problem (Maximise Profit subject to capacity constraint)

The Knapsack Problem	
Explicit Model	
x_1, x_2, x_3, x_4	Quantity of each type of item
maximise	$1000x_1 + 7000x_2 + 3000x_3 + 250x_4$
subject to	$50x_1 + 450x_2 + 100x_3 + 20x_4 \leq 16200$
	$0.4x_1 + x_2 + 2x_3 + 0.1x_4 \leq 32$
	$x_1, x_2, x_3, x_4 \geq 0$
	$x_1, x_2, x_3, x_4 \in \mathbb{Z}$

Decision Variables
Objective Function
Constraints
Weight capacity
Volume capacity
Non-Negativity
Integrality

S	Set of product types
x_i	Number of items from each type
R_i	Revenue from each unit of type i
W_i	Weight of each unit of type i
V_i	Volume of each unit of type i
W_{MAX}	Total weight capacity of the truck
V_{MAX}	Total volume capacity of the truck
maximise	$\sum_{i \in S} R_i x_i$
subject to	$\sum_{i \in S} W_i x_i \leq W_{MAX}$
	$\sum_{i \in S} V_i x_i \leq V_{MAX}$
	$x_i \geq 0 \quad \forall i \in S$
	$x_i \in \mathbb{Z} \quad \forall i \in S$

How to solve mathematical models examples.

Question 1

The shaded area in the figure below represents the feasible region of a linear programming problem. The (unknown) objective function is to be maximised. Decide for each the following statements whether it is true or false. In each case, give an example objective function to justify your answer.

1. The point $(0, 0)$ cannot be an optimal solution.

false: $\text{Min}(x_1 + x_2)$

2. If the objective function has a larger value at point $(3, 3)$ than at points $(0, 2)$ and $(6, 3)$, then $(3, 3)$ must be an optimal solution.

True: $\text{Max}(-x_1 + 6x_2)$

3. If point $(3, 3)$ is an optimal solution and multiple optimal solutions exist, then either $(0, 2)$ or $(6, 3)$ must also be an optimal solution.

(a) True.

$$x_2 = 3$$

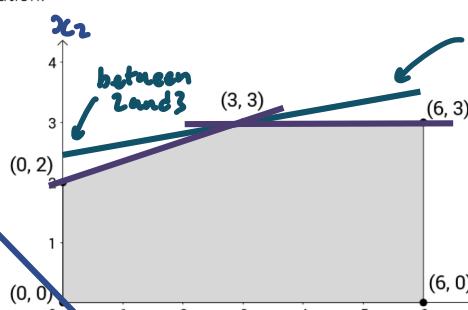
$$\text{Max}(x_1)$$

$$m = \frac{2-0}{3-0} = \frac{1}{3}$$

$$x_2 = \frac{1}{3}x_1 + 1$$

$$3x_2 - x_1 = 3$$

$$\text{Max}(3x_2 - x_1)$$



(b)

to find eqn of this line:

$$m = \frac{3-2-5}{3-0} = \frac{1}{6}$$

$$x_2 = \frac{1}{6}x_1 + 1$$

$$6x_2 - x_1 = 6$$

$$\uparrow$$

keep positive! (not doing like:
 $\frac{1}{6}x_1 - x_2 = -1$)
 cause we be confused.

(a) to find eqn of this line:

$$m = -1$$

$$x_2 = -x_1 + 1$$

$$\text{Min}(x_1 + x_2) \text{ or } \text{Max}(-x_1 - x_2)$$

Question 2

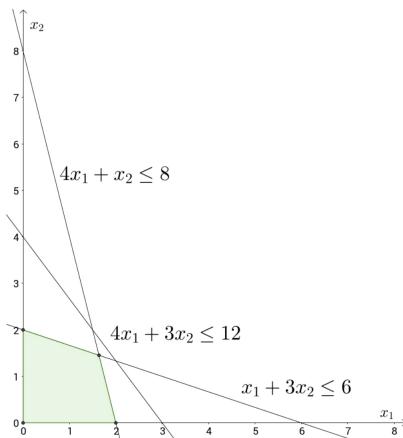
Using the graphical method, determine which of the following three constraints is redundant. Assume that non-negativity conditions for the decision variables x_1 and x_2 also apply.

$$x_1 + 3x_2 \leq 6$$

$$4x_1 + 3x_2 \leq 12$$

$$4x_1 + x_2 \leq 8$$

Solution 2



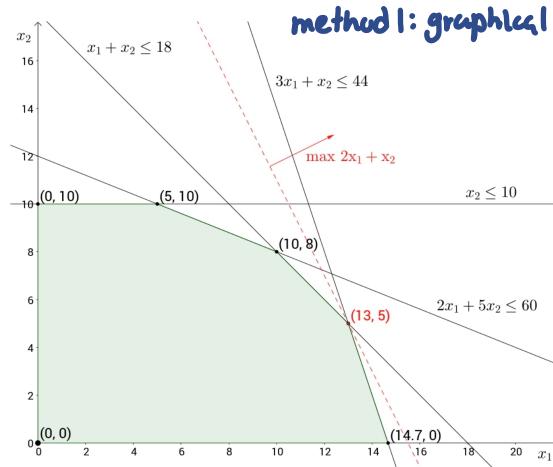
The constraint b) $4x_1 + 3x_2 \leq 12$ is redundant, because removing it would not change the feasible region.

Question 3

Using the graphical method solve the following problem:

$$\begin{array}{ll} \text{maximize} & 2x_1 + x_2 \\ \text{subject to} & x_2 \leq 10 \\ & 2x_1 + 5x_2 \leq 60 \\ & x_1 + x_2 \leq 18 \\ & 3x_1 + x_2 \leq 44 \\ & x_1 \geq 0, x_2 \geq 0 \end{array}$$

Solution 3



method 1: graphical

method 2: corner point evaluation.

Corner point	Objective value
(0, 0)	0
(0, 10)	10
(5, 10)	20
(10, 8)	28
(14.7, 0)	29.3
(13.5, 5)	31

Mathematical Modelling Examples

Question 4 (simple min cost prob. without transit)

A car manufacturer has two factories that produce buses, which are then sold to three local authorities. The production costs are the same at both factories, while the cost of shipping each bus is shown below:

	Local Authority		
	1	2	3
Factory A	£800	£700	£400
B	£600	£800	£500

A total of 60 buses are produced and shipped per week. Each factory can produce and ship up to 50 trucks per week, so there is flexibility on how to divide the production between the two so as to reduce shipping costs. However, each local authority must receive exactly 20 buses per week. We seek to determine how many buses to produce at each factory, and how will they be shipped in order to minimize costs.

Formulate a model for this problem (the explicit form would suffice).

Solution 4.a

Decision variables: $x_{A1}, x_{A2}, x_{A3}, x_{B1}, x_{B2}, x_{B3}$

Objective function: $\min 800x_{A1} + 700x_{A2} + 400x_{A3} + 600x_{B1} + 800x_{B2} + 500x_{B3}$

$$\begin{array}{l} \text{Constraints:} \\ x_{A1} + x_{A2} + x_{A3} \leq 50 \\ x_{B1} + x_{B2} + x_{B3} \leq 50 \\ x_{A1} + x_{B1} = 20 \\ x_{A2} + x_{B2} = 20 \\ x_{A3} + x_{B3} = 20 \\ x_{A1}, x_{A2}, x_{A3}, x_{B1}, x_{B2}, x_{B3} \geq 0 \end{array}$$

Question 5 (min cost prob. without transit)

An airline is expanding and would like to offer direct flights from London to Dublin, Zurich, and Frankfurt. The estimated fares and expected number of passengers and the characteristics of the two types of aircraft available are provided in the tables below.

The airline's objective is to maximise profits by selecting the optimum number of aircrafts from each type that must be deployed, and the number of return flights that must be scheduled for each destination. The airline is very keen to offer at least as many seats as required to meet current levels of demand.

- Formulate an explicit model formulation for this problem.
- Find a solution to this problem using Excel Solver.

Destination	Fare (one way)	Demand outbound	Demand inbound
Dublin	£175	650 p / day	550 p / day
Zurich	£230	450 p / day	500 p / day
Frankfurt	£200	760 p / day	700 p / day

Aircraft type	Seating capacity	Max utilisation	Capital Expenditure	Return flight to Dublin		Return flight to Zurich		Return flight to Frankfurt	
				Duration	OpEx	Duration	OpEx	Duration	OpEx
A320	150	12 h / day	£23,000 / day	3.5 h	£31,000	5.0 h	£47,000	4.5 h	£42,000
A321	230	11 h / day	£32,000 / day	2.5 h	£33,000	4.0 h	£59,000	3.0 h	£51,000

Objective function:

The objective of this problem is to maximise the profit. Where: $\text{Profit} = \text{Revenue} - \text{Cost}$

Destination	Revenue=Fare *demand	Cost	
		Operational Expenditure	Capital Ex. (independent of destination)
Dublin	$175 * (650 + 550)$	$31000 * x_{320}^{DUB} + 33000 * x_{321}^{DUB}$	
Zurich	$230 * (450 + 500)$	$47000 * x_{320}^{ZUR} + 59000 * x_{321}^{ZUR}$	$23000 * n_{320} + 32000 * n_{321}$
Frankfurt	$200 * (760 + 700)$	$42000 * x_{320}^{FRA} + 51000 * x_{321}^{FRA}$	

Explicit model formulation for this problem:

Decision variables: n_{320}, n_{321} (number of aircraft from each class)
 $x_{320}^{DUB}, x_{320}^{ZUR}, x_{320}^{FRA}$ (flights with specific aircraft per destination)
 $x_{321}^{DUB}, x_{321}^{ZUR}, x_{321}^{FRA}$

Objective function:

$$\text{maximize } (175 * (650 + 550) + 230 * (450 + 500) + 200 * (760 + 700)) - (23000 * n_{320} + 32000 * n_{321}) - (31000 * x_{320}^{DUB} + 33000 * x_{321}^{DUB}) - (47000 * x_{320}^{ZUR} + 59000 * x_{321}^{ZUR}) - (42000 * x_{320}^{FRA} + 51000 * x_{321}^{FRA})$$

Constraints:

$$\text{subject to } \begin{aligned} 150 * x_{320}^{DUB} + 230 * x_{321}^{DUB} &\geq 650 \\ 150 * x_{320}^{ZUR} + 230 * x_{321}^{ZUR} &\geq 500 \\ 150 * x_{320}^{FRA} + 230 * x_{321}^{FRA} &\geq 760 \end{aligned}$$

$$\begin{aligned} 3.5 * x_{320}^{DUB} + 5 * x_{320}^{ZUR} + 4.5 * x_{320}^{FRA} &\leq 12 * n_{320} \\ 2.5 * x_{321}^{DUB} + 4 * x_{321}^{ZUR} + 3 * x_{321}^{FRA} &\leq 11 * n_{321} \end{aligned}$$

$$n_{320}, n_{321}, x_{320}^{DUB}, x_{320}^{ZUR}, x_{320}^{FRA}, x_{321}^{DUB}, x_{321}^{ZUR}, x_{321}^{FRA} \geq 0$$

} capacity: passenger demand (usually \geq)
 } capacity: time supply (usually \leq)
 } non-negativity.

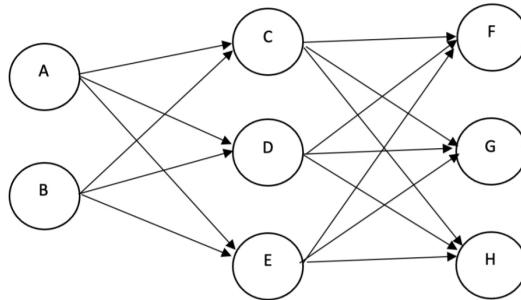
Question 2

A manufacturer has two factories at A and B and seeks to transport goods to their retailers in F, G and H. The following network can be used to represent their supply chain.

(min cost prob. with transit)

- Formulate the problem that minimises the cost of transporting cargo G1 from the supply points to the retailers, ensuring that all demand is satisfied.
- Given the cost table below, solve the above problem using PuLP.
- The manufacturer seeks to extend their product line-up and expects to introduce two additional products into the supply chain (G2 and G3). Re-formulate the problem to incorporate the new commodity types, ensuring link capacities are not exceeded.

(min cost multicommodity prob. with transit)



Node	Demand		
	G1	G2	G3
A	75	25	60
B	75	45	40
C	0	0	0
D	0	0	0
E	0	0	0
F	-50	-30	-50
G	-60	-15	-40
H	-40	-25	-10

Link	Cost	Capacity
AC	5	150
AD	8	150
AE	7	100
BC	9	100
BD	5	150
BE	4	100
CF	3	75
CG	4	90
CH	8	150
DF	7	130
DG	5	100
DH	6	120
EF	9	160
EG	4	90
EH	4	90

Solution 2

a.

Sets	N	Set of nodes
	E	Set of edges
Parameters	c_{ij}	Transportation costs of link (i, j)
	D_i	Supply/Demand of each node i
	L_{ij}	Link capacity
Decision Variables	x_{ij}	Integer: Number of goods transported at each link (i, j)

$$\text{minimise} \sum_{(i,j) \in E} x_{ij} c_{ij}$$

Subject to

$$\sum_{j:(i,j) \in E} x_{ij} - \sum_{l:(j,i) \in E} x_{ji} = D_i \quad \forall i \in N$$

capacity constraints of each arc

$$x_{ij} \leq L_{ij} \quad \forall (i,j) \in E$$

$$x_{ij} \geq 0 \quad \forall (i,j) \in E$$

b. Solution in notebook provided.

c.

$$\text{minimise} \sum_{(i,j) \in E} x_{ij}^l c_{ij}$$

Subject to

$$\sum_{j:(i,j) \in E} x_{ij}^l - \sum_{l:(j,i) \in E} x_{ji}^l = D_i \quad \forall i \in N, l \in C$$

$$\sum_{l \in C} x_{ij}^l \leq L_{ij} \quad \forall (i,j) \in E$$

$$x_{ij}^l \geq 0 \quad \forall (i,j) \in E, l \in C$$

C4 Introduction to Logistics

Introduction to Logistic

Definitions.

- **Logistics vs Supply Chain Management:**

Logistic focus on efficient MOVEMENT and STORAGE of goods.

SCM is a superset of supply, production and distribution

- Potential issues in Logistics:

- Excess costs
- Poor timing (need more storage)
- Upstream disruptions (stagnation in supply)
- Downstream disruptions. (affect information flow)

- Mode of transport:

Air: High value, urgent, perishable goods

Sea/Rail: Bulk, heavy, non-perishable goods

- Why hold inventory (from inventory management)

- lead time (buffer against delay)
- uncertainty (demand fluctuation)
- reliability
- economic of scale (bulk buying)
- speculation.

Inventory Management*

→ I am orderer, how much to order?

1. Economic Order Quantity (EOQ) (used when stock arrives instantly)

- to find optimal order volume, Q^* and reorder frequency, D/Q^* that balances ordering and holding cost.

$$\text{Total Cost} = \text{Ordering Cost} + \text{Holding Cost}$$

$$= \frac{DS}{Q} + \frac{Qh}{2}$$

$$\frac{dT}{dQ} = 0$$

D: Demand Rate

S: Order/Setup Cost per order

Q: Order volume

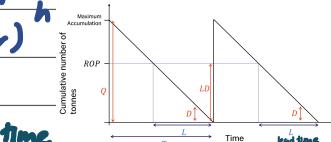
h: Holding cost per unit
($h = I \times C$, I is handling %, C is unit cost, sometimes given h straightaway)

$$(\text{EOQ}) \text{ Optimal Order Volume}, Q^* = \sqrt{\frac{2SD}{h}}$$

• reorder freq = D/Q^* (order per year)

• cycle time = Q^*/D (years)

• reorder point, ROP = $L \times D$ L: lead time



→ I am producer, how much to produce?

2. Economic Production Quantity (EPQ) (used when stock replenished gradually over time)

- to find optimal production volume, Q^* that balances producing and holding costs.

$$\text{Total Cost} = \text{Producing Cost} + \text{Holding Cost}$$

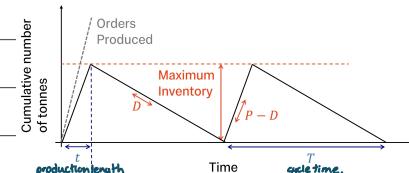
$$= \frac{DS}{Q} + \frac{Qh}{2} (1-x), x = \frac{D}{P}$$

$$(\text{EPQ}) \text{ Optimal Production Volume}, Q^* = \sqrt{\frac{2SD}{h(1-x)}}$$

• Maximum inventory = $Q^* \times (1-x)$

• Production length (cycle time) = Q^*/P

• number of production = D/Q^*



Logistic Costs

Cost of Movement

1. Handling: packaging, loading and unloading

2. Transportation:
cost = $C_f + C_v Q$
(C_f fixed cost: fleet composition, staff schedule)
(C_v var. cost: driver wages, fuel and depreciation)

1. Rent (cost maintaining the facility)

Service cost
storage cost*
(stock management (space, handling and insurance))

2. Wait (losses incurred by delaying pushing the product to the market)

capital cost
(opportunity cost: money tied to unsold inventory)

risk cost
(consequence of theft, deterioration, damage)

Cost of Holding.

EOQ's Examples

Question 2

A company wants to determine their inventory management policy. Their annual consumption is 120,000 units, with order costs rising to £500, unit costs are £50 and handling cost is 10% of unit costs.

$$\text{demand rate, } D = 120000 \text{ unit/year.}$$

- The optimal order volume. $Q^* = \sqrt{2SD/h} = \sqrt{2 \times 120000 \times 500/50} = 4899$
- Number of orders required per year. $D/Q^* = 120000/4899 = 24.5 \text{ per years}$
- Total ordering, total holding costs, and inventory.

(a)

Handling (I) is expressed as a percentage of unit costs, and therefore $I = 0.1$. The unit cost is 50, and therefore the holding costs are: $h = IC = 50 \times 0.1 = 5$

$$\text{Optimal Order Volume: } Q^* = \sqrt{2SD/h} = \sqrt{2 \times 120000 \times 500/5} = 4899$$

(b)

$$\text{Orders per year: } D/Q^* = 120000/4899 = 24.5 \text{ per years}$$

$$\text{Cycle Time: } Q^*/D = 4899/120000 \times 365 = 14.9 \text{ days}$$

(c)

$$\text{Total ordering costs: } O = \text{Orders per year} \times S = 500 \times 24.5 = 12247$$

$$\text{Total holding costs: } H = hQ/2 = 5 \times 4899/2 = 12247$$

$$\text{Inventory costs: } O + H = 12247 + 12247 = 9798$$

(this question need to be careful with units)

Question 3

The demand for product A is 800 units every 2 weeks. The setup cost for placing an order to replenish inventory is £50. Depreciation of the product is accounted by an annual holding cost of £5 per unit. Transportation costs rise to £0.20 per unit with 5-day lead-time, increasing unit costs to £1.5/unit.

*** IMPORTANT**
this is not C
i.e. unit cost,
but this is already
holding cost, h (but unit
is per year)

- Determine the economic order quantity.
- Calculate order cycle.
- Calculate the reorder level.

Solution

(a) Bi-weekly Holding costs: $h = 5 \times 2/52 = 0.192$

$$\text{Optimal Order Volume: } Q^* = \sqrt{2SD/h} = \sqrt{2 \times 50 \times 800/0.192} = 645 \text{ no time unit.}$$

(b) Order Cycle: $Q^*/D = 645/800 = 0.8 \text{ bi - weeks} = 11.3 \text{ days}$

(c) Reorder Point: $ROP = 800/14 \times 5 = 286$ – remember to change demand to daily demand.

unit day⁻¹ day unit

since D is in
bi-weekly unit, convert h
from year unit to bi-weekly unit.

EPQ's Examples

Question 5

A company produces 100,000 units yearly and estimate that their annual demand is 85,000 units.

Production setup costs rise to £500, with holding costs approximating £50. This company operates 270 days a year.

- The optimal production volume.
- Number of production cycles required every year and their length.
- Total ordering, holding costs, and inventory costs.

Solution

$$\text{Optimal Order Volume: } Q^* = \sqrt{2SD/h(1 - \frac{D}{P})} = \sqrt{2 \times 85000 \times 500/50(1 - \frac{85000}{100000})} = 3367$$

$$\text{Production cycles per year: } \frac{D}{Q^*} = \frac{85000}{3367} = \sim 25$$

$$\text{Production Length: } \frac{Q^*}{P} = \frac{3367}{100000} \times 270 = 9.1 \text{ days}$$

(a)

$$\text{Total production costs: } O = S \times \frac{D}{Q^*} = 500 \times \frac{85000}{3367} = 12624$$

$$\text{Total holding costs: } H = hQ(1 - \frac{D}{P})/2 = 50 \times 3367(1 - \frac{85000}{100000})/2 = 12624$$

$$\text{Inventory costs: } O + H = 12624 + 12624 = 25248$$

c5 Freight Transport Network

Freight Transport Network (Move goods efficiently)

Time Series Forecasting.
(Estimate demand, D that we used in c4)

1. Definition

- Level (a) : Baseline where D oscillate.
- Trend (b) : Rate of change .
- Seasonality (S) : Periodic Movement tied to calendar.
- Cyclical Movement (c) : Periodic cycle event not tied to calendar
- Random Fluctuation (e) : Noise

2. Forecasting Methods. *

1. Moving average: Average of last M obs. *

$$\hat{y}_{t+1} = \frac{\sum \text{last } M \text{ observations}}{M}$$

(smooths out noise but lags behind trends)

2. Simple Exponential Smoothing: Used when no trend and no seasonality only level, α

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha) \hat{y}_t$$

(high α makes it responsive but unstable.
low α make it smooth but lags behind)

3. Holt's Model: Adds a trend component, β

4. Winter's Model: Adds a seasonality component, γ (but no trend)

5. Holt-Winter's Model: Adds both trend and seasonality on top of level (α, β, γ)

3. Forecasting Error (pp2022q2b)

1. Error: $E_t = \text{Forecast} - \text{Actual} (F_t - D_t)$

2. Mean Squared Error: $\frac{1}{n} \sum E_t^2$ (MSE) (penalised large error heavily)

3. Mean Absolute Percentage Error: $\frac{100}{n} \sum \left| \frac{E_t}{D_t} \right|$ (MAPE)

4. Mean Absolute Deviation: $\frac{1}{n} \sum |E_t|$ (MAD) (Magnitude of error)

5. Bias: $\sum E_t$ (sum of error)

6. Tracking Signal: $\sum \frac{|E_t|}{\text{MAD}}$

min distance

$\sum_{j \in S} y_j = n$

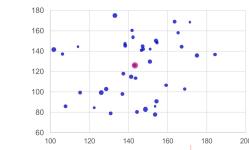
min cost

Facility Location (Placing the Warehouse)

1. continuous problem (center of gravity) → one facility only .

Objective: Minimise the sum of weighted distances between customer locations and the coordinates of the facility (as determined by the problem).

D	Set of customers
x, y	Coordinate values of facility
w_i	Weight of customer i
x_i, y_i	Coordinate of customer i
minimize	$\sum_{i \in D} w_i \sqrt{(x - x_i)^2 + (y - y_i)^2}$



2. Single - Facility Network Location Problem

→ one facility only , and each customer served by one, and infinite capacity

Single-Facility Network Location Problem

In its simplest form, the network facility location problem needs to pick a single depot, among a set of candidates.

D	Set of customers
S	Set of candidate facility locations
x_{ij}	Is customer i served by facility j?
y_j	Is a facility open in location j?
d_{ij}	Distance from customer i and facility j
minimize	$\sum_{i \in D} \sum_{j \in S} x_{ij} d_{ij}$
subject to	$\sum_{j \in S} x_{ij} = 1 \quad \forall i \in D$ $\sum_{i \in D} y_j = 1$ $x_{ij} \leq y_j \quad \forall i \in D, j \in S$ $x_{ij}, y_j \in \{0,1\} \quad \forall i \in D, j \in S$



Every customer i will be served by a single facility
We require one and only one facility j to be open
Facility j is only allowed to serve a customer if it is open
 x_{ij} and y_j are boolean variables

3. Multi-Facility Network Location Problem

→ multiple facility, but each customer served by one and infinite capacity .

4. Capacitated Facility Location Problem

→ flexible no.of facility, each customer can be served by multiple facility , but finite capacity .

Capacitated Facility Location Model – Mathematical Model

minimize	$\sum_{i \in D} \sum_{j \in S} x_{ij} t_{ij} + \sum_{j \in S} c_j y_j$	(1)
subject to	$\sum_{i \in D} x_{ij} = 1 \quad \forall i \in D$	
capacity & logic:	$\sum_{i \in D} x_{ij} \leq P_j y_j \quad \forall j \in S$	(2)
input:	$\begin{cases} x_{ij} = 1 & \text{if facility } j \text{ is open} \\ x_{ij} = 0 & \text{if facility } j \text{ is closed} \\ y_j = 1 & \text{if all firms } (x_{ij}) \text{ must be open} \\ y_j = 0 & \text{if all firms } (x_{ij}) \text{ must be closed} \end{cases}$	(3)
capacity if facility j is open	$\sum_{i \in D} x_{ij} \leq P_j y_j \quad \forall j \in S$	(4)
capacity if facility j is closed	$\sum_{i \in D} x_{ij} \leq n_{\min} \quad \forall j \in S$	(5)
constraint if facility j is open	$x_{ij} \geq 0 \quad \forall i \in D, j \in S$	(6)
constraint if facility j is closed	$y_j \in \{0,1\} \quad \forall i \in D, j \in S$	

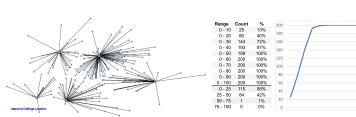
The objective function (1) seeks to minimize the total cost of facility operation and material distribution

- x_{ij} is not boolean, but rather defines the quantity of material flowing from facility j to customer i
- The sum of materials transported to any customer i must be equal to their demand Q_i (2)
- The sum of materials transported from any facility j shall not exceed their total capacity P_j (3)
- If a facility j is not open ($y_j = 0$), then all flows x_{ij} for j shall be equal to 0. (also 3)
- Constraint (4) limits the upper and lower number of open facilities – will be implemented as 2 separate constraints

Level of Service (How much facility we act. need ?)

Level of Service Calculations – 5 Facilities

Facilities	2 Distance	Min. Distance	Max. Distance	Avg. Distance
1	7739	1	78.4	38.7
2	5489	0	27.2	27.2
3	5000	0	53.4	25.0
4	4640	0	56.1	23.2
5	4521	0	56.1	22.6



Time Series Forecasting Examples

Question 1

The demand for a product in each of the last five months is shown below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Demand	5714	6446	7511	7949	8660	9335	9256	10113	11210	11568	11586	

moving average

simple exponential smoothing

forecasting error

- a. Use a four-month moving average to generate a forecast for demand for every month. What is your forecast for month 12?
- b. Apply simple exponential smoothing using a smoothing constant of 0.9 to generate a monthly demand forecast.
- c. Apply simple exponential smoothing using a smoothing constant of 0.8 to generate a monthly demand forecast.
- d. What forecasting method better represents the demand pattern shown? Discuss.

Solution 1

a.

Start forecasting at Month 5, as we are using a 4-month average. We can use an equal weighted forecast:

$$\hat{y}_t = \frac{(y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4})}{4} \quad \text{eg. } \hat{y}_5 = \frac{y_1 + y_2 + y_3 + y_4}{4} = \frac{5714 + 6446 + 7511 + 7949}{4} = 6905$$

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Demand					6905	7642	8364	8800	9341	9979	10537	11119

Students are encouraged to test different weighting factors.

b.

Start forecasting at Month 2 using the following equation:

$$\hat{y}_t = 0.9y_{t-1} + 0.1\hat{y}_{t-1}$$

$$\hat{y}_2 = 0.9y_1 + 0.1\hat{y}_1$$

$$= y_1$$

$$= 5714$$

$$\hat{y}_3 = 0.9y_2 + 0.1\hat{y}_2$$

$$= 0.9(6446) + 0.1(5714)$$

$$= 6377.8$$

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Demand		5714	6373	7397	7894	8583	9260	9256	10028	11092	11520	11579

c.

Start forecasting at Month 2 using the following equation:

$$\hat{y}_t = 0.8y_{t-1} + 0.2\hat{y}_{t-1}$$

$$\hat{y}_3 = 0.8y_2 + 0.2\hat{y}_2$$

$$= 0.8(6446) + 0.2$$

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Demand		5714	6300	7269	7813	8490	9166	9238	9938	10955	11445	11558

d.

For the period between April and November, the performance indicators yield:

Method	MSE	MAPE	MAD	BIAS	TS
Moving Averages	2226704	14.3	1452	-10161	-7
Exponential (0.9)	565105	7.4	653	-6517	-10
Exponential (0.8)	668244	8.2	730	-7304	-10

(not really recommend to revise due to complication)

Facility Location Problem Examples

instead, recommended to revise 2023 q2c

(this example is slightly too complicated cause of the parameter z!)

	AA	BB	CC	DD	Yearly trips
A	£150	£250	£200	£25	40
B	£150	£250	£200	£75	35
C	£75	£200	£150	£125	100
D	£150	£200	£125	£125	25
E	£100	£200	£125	£150	40
F	£175	£175	£125	£125	25
G	£150	£175	£100	£150	50
H	£150	£150	£100	£200	30
Fixed Cost	£43935	£54892	£57820	£31471	-

A consulting company is seeking to establish its new headquarters. They have selected four possible locations AA, BB, CC, and DD, which would each incur a fixed cost described in the above table. Their clients are in 8 regions (A-H).

- Select the optimal location for a single facility that minimises the total cost.
- Formulate the multiple-facility location problem that minimises the total cost, considering that each consultant can take 10 trips at most annually.
- Solve the above problem using PuLP, using the accompanying notebooks provided.
- How would you modify your above formulation if:
 - Only 20 consultants can be assigned to each office.
 - Each customer can only be serviced by one office.
 - Both of the above conditions apply.
- Solve the mathematical problem described in the above sub-question. Does the answer change substantially from sub-question c)?

- Calculate the total cost for facility location using the following equation:

$$\text{Fixed Costs} + \sum_{i \in F} (\text{yearly trips}) \times (\text{travel cost}) = \text{Total Yearly Cost}$$

The cost of each facility is therefore:

	AA	BB	CC	DD
Total Cost	86810	124267	107070	73346

DD is the optimal location.

- The formulation is:

Sets	S	Set of suppliers
	D	Set of customers
Parameters	c_j	Fixed cost of establishing an office at site j
	t_{ij}	Travel cost between each site j and customer i
	Q_i	Demand of each customer i
	P	Consultant trip capacity
	M	Big number
Decision Variables	x_{ij}	Integer: Number of trips allocated from each site j to each client i
	y_j	Boolean: Whether office is established at each site j
	z_j	Integer: Number of consultants hired at each office

$$\text{minimise} \sum_{i \in D} \sum_{j \in S} x_{ij} t_{ij} + \sum_{j \in S} c_j y_j$$

Subject to

$$\sum_{j \in S} x_{ij} = Q_i \quad \forall i \in D$$

$$\sum_{j \in S} y_j \geq 1$$

$$y_j M \geq z_j \quad \forall j \in S$$

$$\sum_{i \in D} x_{ij} \leq P z_j \quad \forall j \in S$$

$$x_{ij} \geq 0 \quad \forall i \in D, \forall j \in S$$

$$y_j \in \{0,1\} \quad \forall j \in S$$

$$z_j \geq 0 \quad \forall j \in S$$

Capacitated Facility Location Model – Mathematical Model
$\begin{aligned} & \text{minimize} && \sum_{i \in D} \sum_{j \in S} x_{ij} t_{ij} + \sum_{j \in S} c_j y_j \quad (1) \\ & \text{subject to} && \sum_{j \in S} x_{ij} = Q_i, \quad \forall i \in D \quad (2) \\ & && \sum_{i \in D} x_{ij} \leq P y_j, \quad \forall j \in S \quad (3) \\ & && u_{\text{max}} \leq \sum_{j \in S} y_j \leq u_{\text{max}} \quad (4) \\ & && x_{ij} \geq 0, \quad \forall i \in D, j \in S \quad (5) \\ & && y_j \in \{0,1\}, \quad \forall i \in D, j \in S \quad (6) \end{aligned}$

d. i) $z_j \leq 20 \quad \forall j \in S$

ii) Create a new variable that indicates if office is assigned to demand $v_{ij} \in \{0,1\}$

$$\sum_{j \in S} v_{ij} = 1 \quad \forall i \in D$$

$$x_{ij} \leq v_{ij}M \quad \forall i \in D, \forall j \in S$$

Question 2 (also too complicated for revision, due to type of parameter)

A manufacturer intends to open a business in a new country. Over the next year, they expect to sell 250,000 units in the North, 140,000 in the West, 180,000 in the South, and 110,000 in the East. Three types of factories can be built: Small, Medium, and Large, which can produce 200,000, 300,000 and 400,000 units of material every year. These can be built at four locations: A, B, C, and D.

- a) Formulate the facility location problem that minimises the cost of satisfying the demand.
- b) Given the table of costings below, solve the above problem using PuLP.

	A	B	C	D
Annual Fix Cost Small (£ million)	6	5.5	5.6	6.1
Annual Fix Cost Medium (£ million)	8.5	7.8	7.9	8.7
Annual Fix Cost Large (£ million)	10	9.2	9.3	10.2
North	£211	£232	£238	£299
West	£232	£212	£230	£280
South	£240	£230	£215	£270
East	£300	£280	£270	£225

Solution 2

a.

Sets	S	Set of suppliers
	D	Set of customers
	F	Set of factory types
Parameters	c_{jf}	Fixed cost of establishing a factory of type t at site j
	t_{ij}	Transportation cost between each site j and customer i
	Q_i	Demand of each customer i
	P_f	Production capacity of factory type f
Decision Variables	x_{ij}	Integer: Number of goods allocated from each site j to each client i
	y_{jf}	Integer: Number of factories of type t are established at each site j

$$\text{minimise} \sum_{i \in D} \sum_{j \in S} x_{ij} t_{ij} + \sum_{j \in S} \sum_{f \in F} c_{jf} y_{jf}$$

Subject to

$$\sum_{j \in S} x_{ij} = Q_i \quad \forall i \in D$$

$$\sum_{j \in S} \sum_{f \in F} y_{jf} \geq 1$$

$$\sum_{i \in D} x_{ij} \leq \sum_{f \in F} P_f y_{jf} \quad \forall j \in S$$

$$x_{ij} \geq 0 \quad \forall i \in D, \forall j \in S$$

$$y_{jf} \geq 0 \quad \forall j \in S, \forall f \in F$$

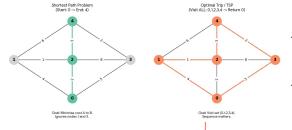
Capacitated Facility Location Model – Mathematical Model	
D	Set of customers
S	Set of candidate facility locations
x_{ij}	Quantity supplied to i served by j
y_j	Is a facility open in location j ?
t_{ij}	Unit shipping cost from j to i
p_{\min}	Minimum number of open locations
p_{\max}	Maximum number of open locations
c_f	Facility f operating cost (if open)
Q_i	Material demand by customer i
P_f	Total supply capacity in facility f
$\text{minimise} \quad \sum_{i \in D} \sum_{j \in S} x_{ij} t_{ij} + \sum_{j \in S} c_f y_j \quad (1)$	
$\text{subject to} \quad \sum_{j \in S} x_{ij} = Q_i \quad \forall i \in D \quad (2)$	
$\sum_{j \in S} y_j = p_f \quad \forall f \in F \quad (3)$	
$n_{\min} \leq \sum_{j \in S} y_j \leq n_{\max} \quad (4)$	
$x_{ij} \geq 0 \quad \forall i \in D, j \in S \quad (5)$	
$y_j \in \{0,1\} \quad \forall i \in D, j \in S \quad (6)$	

c6. Touring and Routing

Touring and Routing

Touring

(Single vehicle, visit all nodes)



} touring & shortest path prob.
} touring visit all nodes
and shortest path doesn't
but both tries to save costs.

$$\text{no. of tour} = (n-1)!/2$$

1. Travelling Salesman Problem (TSP) (one vehicle, infinite capacity)

N	Set of vertices (cities)
x_{ij}	Are vertices i and j connected?
c_{ij}	Cost of travel from i to j
minimize	$\sum_{i \in N} \sum_{j \in N, i \neq j} x_{ij} c_{ij}$ minimise total cost
subject to	$\sum_{j \in N, i \neq j} x_{ij} = 1 \quad \forall i \in N$ must leave every node once $\sum_{i \in N, i \neq j} x_{ij} = 1 \quad \forall j \in N$ $x_{ij} \in \{0,1\} \quad \forall i, j \in N$



all these are subtour

to only have one route, we have to
get rid of subtour by adding constraints:

method 1: lazy constraints:

$$x_{ij} + x_{ji} \leq 1 \quad \forall i, j \in N$$

n=1 get rid of two nodes subtour
n=2 get rid of three nodes subtour...

method 2: MTZ constraints:

$$u_i - u_j \leq U(1 - x_{ij}) - 1 \quad \forall \begin{cases} i, j \in N \\ i \neq j \\ j \neq 0 \end{cases}$$

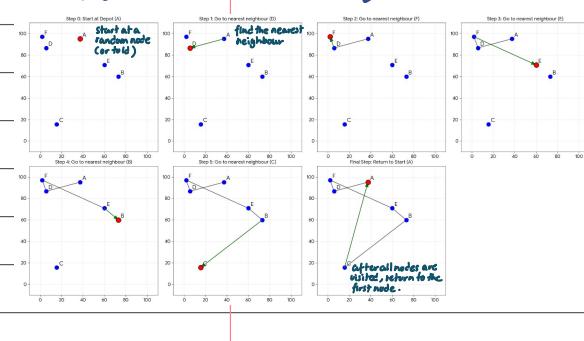
$$x_{ij} \in \{0,1\} \quad \forall i, j \in N$$

$$u_i \leq U-1 \quad \forall i \in N$$

heuristic

2. Nearest Neighbour Algorithm

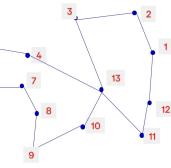
(Greedy Heuristic Approach, i.e. do whatever seems to be the best choice)



1. Multiple Travelling Salesman Problem (mTSP) (m vehicles, infinite capacity)

Multiple Travelling Salesman Problem

N	Set of vertices (cities)
x_{ij}	Are vertices i and j connected?
c_{ij}	Cost of travel from i to j
d	Index of depot vertex
m	Number of tours from d
minimize	$\sum_{i \in N} \sum_{j \in N, i \neq j} x_{ij} c_{ij}$ minimise total cost
subject to	$\sum_{j \in N, i \neq j} x_{ij} = m \quad \text{Exactly } m \text{ arcs (and tours) commence from vertex } d$ $\sum_{i \in N, i \neq d} x_{id} = m \quad \text{Exactly } m \text{ arcs (and tours) terminate at vertex } d$ $\sum_{j \in N, i \neq d} x_{ij} = 1 \quad \forall i \in N, i \neq d \quad \text{As in TSP, but with the exception of } d$ $\sum_{i \in N, i \neq d} x_{ij} = 1 \quad \forall j \in N, i \neq d \quad \text{As in TSP, but with the exception of } d$ $x_{ij} \in \{0,1\} \quad \forall i, j \in N$ + subtour elimination constraints

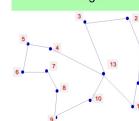


2. Capacitated Vehicle Routing Problem (CVRP) *

(m vehicles, but finite capacity, i.e. payload p)

Capacitated Vehicle Routing Problem

V	Set of vertices
S, D	Set of suppliers, customers $S, D \subset V$
x_{ij}	Are vehicles travelling from i to j ?
f_{ij}	How much cargo to carry from i to j
d_{ij}	Distance between vertices i and j
p	Vehicle payload
Q_i	Demand by customer i
c	Investment cost per vehicle
B	Budget
minimize	$\sum_{i \in N} \sum_{j \in N, i \neq j} x_{ij} d_{ij}$ minimise total distance
subject to	$\sum_{j \in V} f_{ij} - \sum_{j \in V} x_{ij} = 0 \quad \forall i \in D$ $\sum_{i \in V} x_{ij} - \sum_{i \in V} f_{ij} = Q_i \quad \forall i \in S$ $* Q_i \leq f_{ij} \leq (p - Q_i) \quad \forall i, j \in V$ $\sum_{i \in V} x_{ij} \leq p \quad \forall j \in V$ $\sum_{i \in V} x_{ij} \leq 1 \quad \forall i \in S, j \in V$ $\sum_{j \in V} c x_{ij} \leq B \quad \forall i \in S$ $f_{ij} \geq 0 \quad \forall i \in V, \forall j \in V$ $x_{ij} \in \{0,1\} \quad \forall i \in V, \forall j \in V$

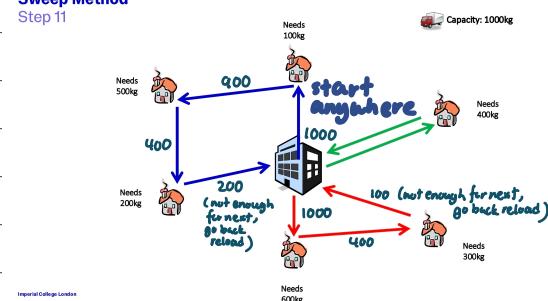


heuristic

3. Heuristic Method 1: The sweep method

Sweep Method

Step 11



4. Heuristic Method 2: The savings method. *

The Savings Method

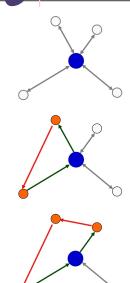
- Developed by Clarke and Wright (1963) – also an iterative approach.
- Initially considers one tour per customer (depot → customer → depot).
- Assuming that vertex O is the depot, the total distance of all tours is:

$$\sum_{j=1..n} (c_{0j} + c_{j0})$$

- Trip pairings are considered to reduce costs (savings):

$$s_{ij} = c_{ij} + c_{oj} - c_{ij}$$

- At each iteration we consider the largest known saving and add it to a tour whose capacity is not violated.
- Algorithm terminates when all potential savings have been evaluated.



The saving methods in action:

Savings Method

Example Problem

Distances

	0	1	2	3	4
0	-	20	25	23	26
1	-	-	20	32	27
2	-	-	-	21	22
3	-	-	-	-	12
4	-	-	-	-	-

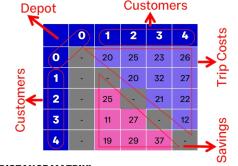
Demands

	Demands
1	2
2	14



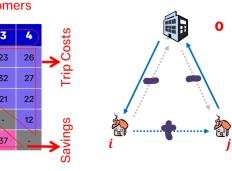
Savings Method

Step 1: Create Savings Table



NOT A DISTANCE MATRIX

- Trip costs: Only on upper triangle, refer to bidirectional travel costs between any two nodes i and j .
- Each initial default trip between customer i and depot 0 will be equal to $c_{0i} + c_{i0}$. In the savings table, both are equal, i.e. $2 \times c_{0i}$
- Savings: Only on lower triangle, savings by combining any 2 customers



Savings Method

Step 2: Rank Savings

	0	1	2	3	4
0	-	20	25	23	26
1	-	-	20	32	27
2	-	-	-	21	22
3	-	-	-	-	12
4	-	-	-	-	-

	Rank	Link	Saving
1	3 \leftrightarrow 4	37	
2	2 \leftrightarrow 4	29	
3	2 \leftrightarrow 3	27	
4	2 \leftrightarrow 1	25	
5	1 \leftrightarrow 4	19	
6	1 \leftrightarrow 3	11	

$$s_{ij} = c_{0i} + c_{0j} - c_{ij} = 20 + 25 - 20 = 25$$

$$s_{ij} = c_{0i} + c_{0j} - c_{ij} = 20 + 23 - 32 = 11 \text{ etc.}$$

From the savings table we obtain and rank all savings that may stem from the combination of trips between any two nodes.

Savings Method

Iteration 1

Tour	Cargo Left
0 \rightarrow 1 \rightarrow 0	15
0 \rightarrow 2 \rightarrow 0	6
0 \rightarrow 3 \rightarrow 0	8
0 \rightarrow 4 \rightarrow 0	13

- We investigate saving #1 ($3 \leftrightarrow 4$)
- If applied to $0 \rightarrow 3 \rightarrow 0$, this becomes $0 \rightarrow 3 \rightarrow 4 \rightarrow 0$
- If applied to $0 \rightarrow 4 \rightarrow 0$, this becomes $0 \rightarrow 4 \rightarrow 3 \rightarrow 0$
- Both equivalent. Let's go ahead with first case.
- Capacity check: New tour is possible since current tour has 8t of cargo left, while customer 4 needs only 7t.

Checking if the "saving" is alright or not doesn't depend on the cost. It depends on the capacity of the cargo vehicle.

Rank	Link	Saving
1	3 \leftrightarrow 4	37
2	2 \leftrightarrow 4	29
3	2 \leftrightarrow 3	27
4	2 \leftrightarrow 1	25
5	1 \leftrightarrow 4	19
6	1 \leftrightarrow 3	11

	Demands
1	2
2	14

	Capacity: 20 units

Savings Method

Iteration 4

Current tours:

Tour	Cargo Left
0 \rightarrow 1 \rightarrow 0	15
0 \rightarrow 2 \rightarrow 0	6
0 \rightarrow 3 \rightarrow 0	1

- We investigate saving #4 ($2 \leftrightarrow 1$)
- Cannot add to $0 \rightarrow 3 \rightarrow 4 \rightarrow 0$ (does not contain 2 or 1)
- Can add to either $0 \rightarrow 2 \rightarrow 0$ or $0 \rightarrow 1 \rightarrow 0$
- Both equivalent. Let's go ahead with first case.
- Capacity check for $0 \rightarrow 2 \rightarrow 1 \rightarrow 0$: OK!

Rank	Link	Saving
1	3 \leftrightarrow 4	37
2	2 \leftrightarrow 4	29
3	2 \leftrightarrow 3	27
4	2 \leftrightarrow 1	25
5	1 \leftrightarrow 4	19
6	1 \leftrightarrow 3	11

	Demands
1	2
2	14

	Capacity: 20 units

Savings Method

Iteration 2

Current tours:

Tour	Cargo Left
0 \rightarrow 1 \rightarrow 0	15
0 \rightarrow 2 \rightarrow 0	6
0 \rightarrow 3 \rightarrow 0	1

- We investigate saving #2 ($2 \leftrightarrow 4$)
- If we add to $0 \rightarrow 3 \rightarrow 4 \rightarrow 0$, this $0 \rightarrow 3 \rightarrow 4 \rightarrow 2 \rightarrow 0$ X
- Capacity check: Not possible, since not enough cargo left in tour to also service customer 2.
- Cannot add to $0 \rightarrow 2 \rightarrow 0 \rightarrow 0$ (customer 4 in other tour)
- Cannot add to $0 \rightarrow 1 \rightarrow 0$ (does not contain 2 or 4)
- We therefore dismiss saving #2
- We investigate saving #3 ($2 \leftrightarrow 3$)
- Cannot add to $0 \rightarrow 3 \rightarrow 4 \rightarrow 0$ (well, not enough cargo after all)
- Cannot add to $0 \rightarrow 2 \rightarrow 0 \rightarrow 0$ (customer 3 in other tour)
- Cannot add to $0 \rightarrow 1 \rightarrow 0$ (does not contain 2 or 3)
- We therefore dismiss saving #3
- No more potential savings left to investigate. We therefore finalize both tours, even though they have some cargo left.

final tour: total savings: $37 + 25 = 62$

Tour: Cargo Left
0 \rightarrow 1 \rightarrow 0 1
0 \rightarrow 3 \rightarrow 4 \rightarrow 0 1

need 7
0 \rightarrow 2 \rightarrow 1 \rightarrow 4 \rightarrow 3 \rightarrow 0 0
0 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 0 0
need 12
0 \rightarrow 2 \rightarrow 1 \rightarrow 4 \rightarrow 0 0

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

X

Touring Example : (Heuristic)

Question 3

Apply the nearest neighbour heuristic once to obtain a travelling salesman route among a set of nodes described with the distance matrix below. Comment on the optimality of the result.

	1	2	3	4	5	6	7
1. 1	0.00	72.01	9.43	91.92	88.51	76.69	106.51
3. 2	72.01	0.00	64.26	61.77	18.87	65.86	69.26
2. 3	9.43	64.26	0.00	82.64	80.13	67.90	97.17
5. 4	91.92	61.77	82.64	0.00	55.17	21.47	15.00
4. 5	88.51	18.87	80.13	55.17	0.00	65.31	58.83
7. 6	76.69	65.86	67.90	21.47	65.31	0.00	35.51
6. 7	106.51	69.26	97.17	15.00	58.83	35.51	0.00

Solution 3

Assuming that we commence our tour from Node 1, the route will be 1->3->2->5->4->7->6->1, with a total cost of 275.03. The Nearest Neighbour heuristic provides an approximate result with no guarantees of optimality. Furthermore, different results may be obtained depending upon the node chosen to depart from initially.

Choose initial centre points P1 = A = (2, 10) and P2 = C = (8, 4).

Routing Example : (Heuristic)

Question 1

The London Assembly is trailing a new recycling program, and has made provisional arrangements for refuse paper to be processed in four different plants around London, all owned by the same company. Their weekly processing capacity is as follows:

Facility	Capacity (tons)
Plant 1	90
Plant 2	50
Plant 3	80
Plant 4	60

Three London boroughs are participating in the pilot study, with the following estimated weekly volumes:

Borough	Supply (tons)
A. Kensington and Chelsea	120
B. Westminster	100
C. Greenwich	110

Aggregate transportation costs (distance, time) per ton between the boroughs and the plants are listed as follows:

		To		
		1	2	3
From	A	14	9	16
	B	11	8	7
C	16	12	10	22

- a. Is this trial viable?
- b. How should collected refuse be distributed to plants to minimize costs using the current plant configurations? Formulate a mathematical model.
- c. Implement the above model using PuLP to obtain a solution.
- d. The recycling company wishes to expand plant capacities to process all demand from the boroughs. Assuming that the additional operating costs per introduced ton of weekly capacities for each plant i is g_i , present an updated formulation that determines the cheapest combined plant upgrade and refuse distribution strategy.

Solution 2

The distance/savings matrix is as follows.

	1	2	3	4	5	6	7	8	9	10
1		25	43	57	43	61	29	41	48	71
2				29	34	43	68	49	66	72
3			39		52	72	96	72	81	89
4				48	48		45	71	71	99
5					25	14	55		27	36
6						18	8	47	77	
7							5	0	15	36
8								19	36	39
9									1	26
10										5
								0	20	49
									86	57
									66	83

We construct a ranked savings table, as shown in the matrix below. The capacity of each vehicle is 23, and we assume that no other constraints apply.

Link	Savings	Link	Savings	Link	Savings
(6, 10)	86	(6, 9)	47	(3, 5)	14
(9, 10)	83	(7, 9)	46	(3, 6)	8
(8, 9)	78	(2, 3)	39	(4, 9)	6
(5, 6)	77	(7, 8)	39	(2, 7)	5
(8, 10)	66	(5, 7)	36	(2, 10)	5
(7, 10)	57	(6, 8)	36	(3, 8)	3
(4, 5)	55	(5, 9)	26	(4, 8)	3
(6, 7)	50	(2, 5)	25	(3, 9)	2
(5, 10)	49	(4, 10)	20	(2, 9)	1
(3, 4)	48	(5, 8)	19	(2, 8)	0
(2, 4)	48	(2, 6)	18	(3, 7)	0
(4, 6)	47	(4, 7)	15	(3, 10)	0

We proceed in steps, while progressing through the ranked savings list.

- Step 1: Add link (6, 10) to I Current tour: I {1, 6, 10, 1}
- Step 3: Add link (9, 1) to I Current tour: I {1, 6, 10, 9, 1}
- Step 4: Add link (8, 9) to I Current tour: I {1, 6, 10, 9, 8, 1}
- Step 5: Skip link (5, 6) (it would violate capacity constraints in tour I)
- Step 6: Skip link (8, 10) (nodes 8 and 10 already exist in tour I)
- Step 7: Skip link (7, 10) (node 10 already exists in tour I)
- Step 9: Link (4, 5) contains two nodes that do not exist in tour I, so we initiate tour II
- Step 10: Add link (4, 5) to II Current tour: II {1, 4, 5, 1}
- Step 11: Add link (6, 7) to I Current tour: I {1, 7, 6, 10, 9, 8, 1}
- Step 12: Tour I now has a load of 23, and therefore we close it.
- Step 13: Skip link (5, 10) (nodes 5 and 10 already exist in our tours)
- Step 14: Add link (3, 4) to II Current tour: II {1, 3, 4, 5, 1}
- Step 15: Skip link (2, 4) (node 4 already exists in tour II)
- Step 16: Skip link (4, 6) (nodes 4 and 6 already exist in our tours)
- Step 17: Skip link (6, 9) (nodes 9 and 6 already exist in our tours)
- Step 18: Skip link (7, 9) (nodes 9 and 7 already exist in our tours)
- Step 19: Add link (2, 3) to II Current tour: II {1, 2, 3, 4, 5, 1}
- We have now included all links to our tours, so we can now close tour II as well. Their combined length (total distance travelled) is 397 units.

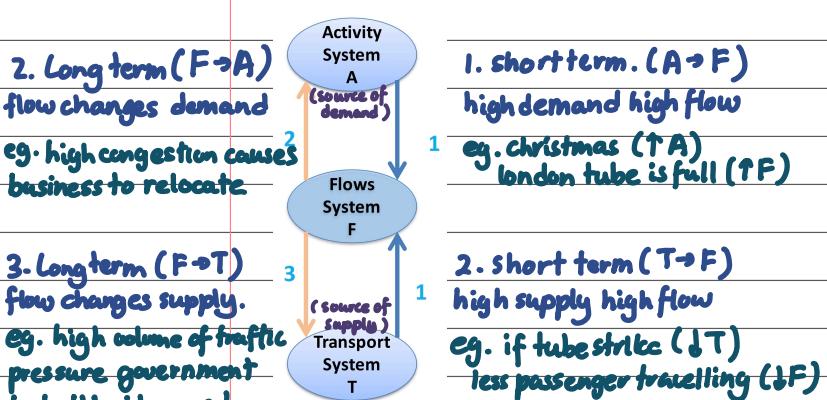
C7. Transport Planning

- Aim: To solve the "Prediction Problem"
- How to predict the impact of an intervention?
 - How to evaluate those impacts are desirable compared to the costs?
 - (eg: Should we widen the road? reduced time travel vs. construction cost)

Transport Planning

Manheim's Analysis Framework

- Activity (A): Landuse, socio-economic factors, where people live or work.
- Transport system (T): Physical networks, vehicles, schedules, prices.
- Flow system (F): Actual traffic, passenger volumes, travel time (resulting from A and T)



long term because business can't relocate immediately, road can't be built overnight.

Congestion and Cost

Private Cost for ONE Traveller:

$$c(q)$$

Total Social Cost for ALL Traveller:

$$q \cdot c(q)$$

Marginal social cost (change in total social cost when you add one more traveller)

$$\begin{aligned} MSC &= \frac{d}{dq}(q \cdot c(q)) \\ &= c(q) + q \cdot \frac{dc}{dq} \end{aligned}$$

product rule.

private cost public cost (externality)

(you pay your own delay)

(everyone else has to pay for the slight increase in delay you caused for being there)

this term is the reason why congestion is a prob. because $\frac{dc}{dq} > 0$ (costs always rises with traffic) $\rightarrow q \cdot \frac{dc}{dq} > 0$
 \rightarrow social cost is always higher than private cost.
 $(C_{at opt_p} < C_{at opt_s})$

Mathematical Formulation and Equilibrium. (Manheim became mathematic's equation)

1. Short term - relationship 1 ($A \rightarrow F, T \rightarrow F$)

$$V = D(A, S); \quad S = J(V, T)$$

volume of travel demanded
activity system attributes
(population, (travel time, income))

transport system service
(capacity, route)

note that you don't see F in the eqn because F is a vector of flow, i.e. $F = (V, S)$

2. Long term - relationship 2 and 3 ($F \rightarrow A, F \rightarrow T$)

$$A = H(T, F);$$

$$T = I(A, F)$$

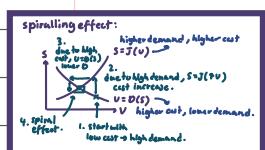
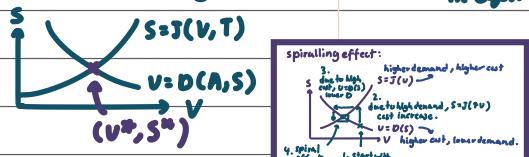
(not really important)

3. Demand-Performance Equilibrium *

A system is said to be equilibrium when volume of travel demanded, V matches the system's ability to supply that travel at a specific service level, S :

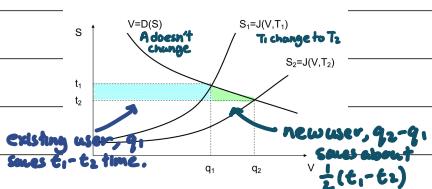
$$V^* = D(A, S^*); \quad S^* = J(V^*, T)$$

flow when system is in equil.



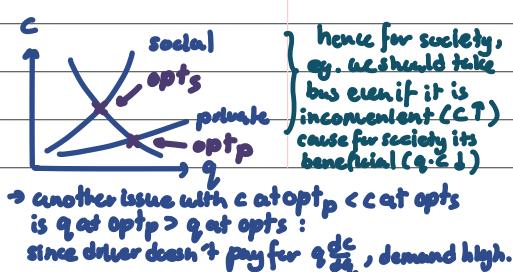
4. Economic Evaluation and "Rule of a Half"

This answers our "Prediction Problem" i.e. if impact of an intervention are desirable compared to the cost.



$$\text{total benefit (both existing and new)} = \frac{1}{2}(q_1 + q_2)(t_1 - t_2)$$

we will check if this is greater than a threshold:
 $\frac{1}{2}(q_1 + q_2)(t_1 - t_2) > \lambda$



Demand-Performance Equilibrium Examples

INTRODUCTION

In a congestible facility, the generalised cost of travel c depends on the level of demand q . Likewise, the level of demand will depend on the generalised cost. Hence

$$\text{Performance: } c = C(q) \rightarrow S = J(V)$$

$$\text{Demand: } q = Q(c) \rightarrow V = D(S)$$

The natural approach to predicting the actual demand and cost that will be experienced by travellers on the network is to seek the simultaneous solution of these equations:

$$c^* = C(q^*)$$

$$q^* = Q(c^*)$$

The pair (q^*, c^*) is called an equilibrium flow pattern (see Manheim's paper).

In general, for large networks or complex patterns of demand, the computation of such equilibria is rather difficult. However, for simple networks, the computations can be straightforward. Consider for example, the simple single link network shown below:



Suppose that the demand for travel on this network from the origin O to the destination D is given by the following (linear) demand model:

$$q = \alpha - \beta c$$

and that the generalised cost of travel from O to D is given by the following (linear) performance model:

$$c = \mu + \lambda q$$

where the quantities $\alpha, \beta, \mu, \lambda$ are all positive parameters.

Tasks

1. Discuss the physical interpretation of the parameters .
2. Compute the equilibrium values q^* and c^* , in terms of these parameters and discuss your result.
3. How do the equilibrium values q^* and c^* change if the network becomes less congestible? *increase c-intercept*
- *HARD** 4. Suppose that the network operator wanted to introduce a fixed toll p for travellers using this facility and wants to know what level of toll would maximise the total revenue generated by the network. How would you extend the model above to answer this question?

*very typical
demand-performance
questions.*

*interesting
and important
to understand .*

$$1) \quad \underbrace{V = D(S)}$$

The demand model is

$$\underbrace{S = J(V)}_{\text{Performance Model}} \quad q = \alpha - \beta c \quad \alpha, \beta > 0 \quad (1)$$

and the performance model is

$$c = \mu + \lambda q \quad \mu, \lambda > 0 \quad (2)$$

α = maximum demand possible

β = sensitivity of demand to changes in generalised cost

μ = minimum generalised cost (e.g. in uncongested conditions)

λ = sensitivity of generalised cost to changes in demand - the congestability of the network.

2)

The equilibrium values of demand and generalised cost q^* and c^* satisfies:

$$q^* = \alpha - \beta c^* \quad (3)$$

$$c^* = \mu + \lambda q^* \quad (4)$$

to find equilibrium solution to demand-performance model

→ solve simultaneous eqn.

$$\Rightarrow q^* = \frac{\alpha - \beta\mu}{1 + \beta\lambda} \quad c^* = \frac{\mu + \lambda\alpha}{1 + \beta\lambda} \quad \text{if } (\alpha > \beta\mu) \quad (5)$$

3)

The equilibrium can be represented as the intersection of the curves representing the demand and generalised cost models **less congestable doesn't mean q reduces. It means c (travel time) doesn't increase significantly as q increases.**

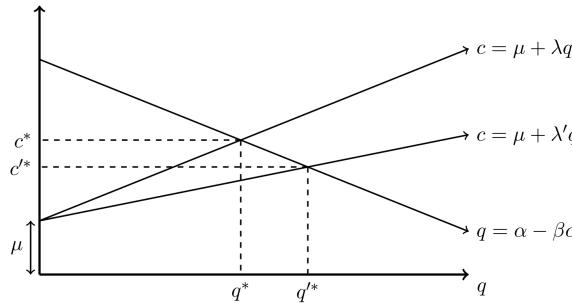
If the network becomes less congestable, then the value of λ reduces, say to λ' , and the equilibrium shift to (c'^*, q'^*) from (c^*, q^*) where $c'^* < c^*$ and $q'^* > q^*$.

Algebraically:

this basically means as λ increases ($\lambda' \rightarrow \lambda$) q^* decreases ($q^* \rightarrow q'^*$) and c^* increases ($c^* \rightarrow c'^*$)

$$\frac{\partial q^*}{\partial \lambda} = \frac{-\beta(\alpha - \beta\mu)}{(1 + \beta\lambda)^2} < 0 \quad (6)$$

$$\frac{\partial c^*}{\partial \lambda} = \frac{\alpha - \beta\mu}{(1 + \beta\lambda)^2} > 0 \quad (7)$$



1. toll relates directly to S (system service level)

4) (quite hard, especially the concept of 2. since c is about generalised cost in time units, we need to convert p's unit).

Suppose that a toll of p is charged. This has the effect of increasing the minimum generalised cost of travel from μ to $(\mu + \theta p)$, where θ is a parameter that converts units of toll (money) into units of travel time (assuming that we are measuring generalised cost in time units). θ is a parameter that characterises travel demand.

Therefore, the equilibrium with the toll p is:

$$q^* = \frac{\alpha - \beta(\mu + \theta p)}{1 + \beta\lambda} \quad c^* = \frac{(\mu + \theta p) + \lambda\alpha}{1 + \beta\lambda} \quad (8)$$

The revenue (in units of money) that the toll raises, is $q^* \times p$

$$R = q^* \times p = \frac{p(\alpha - \beta(\mu + \theta p))}{1 + \beta\lambda} \quad (9)$$

The toll revenue is maximised when $\frac{\partial R}{\partial p} = 0$ and $\frac{\partial^2 R}{\partial p^2} < 0$

$$\frac{\partial R}{\partial p} = \frac{\alpha - \beta\mu - 2p\beta\theta}{1 + \beta\lambda} \quad \frac{\partial^2 R}{\partial p^2} = \frac{-2\beta\theta}{1 + \beta\lambda} \quad (10)$$

So, the optimised toll is:

$$p = \frac{1}{2\beta\theta}(\alpha - \beta\mu) \quad \text{obtained by } \frac{\partial R}{\partial p} = 0 \quad (11)$$

and the maximum revenue is:

$$R = \frac{1}{4\beta\theta} \left[\frac{(\alpha - \beta\mu)^2}{1 + \beta\lambda} \right] \quad (12)$$

c8. Travel Demand Models

The 4-stage model can then be expressed as follows:

$$T_{ijmr} = T_j \frac{r_j}{T_j} \cdot \frac{T_{jm}}{T_j} \cdot \frac{T_{jmr}}{T_{jm}}$$

or

$$T_{ijmr} = f_g(X_i, c_i) \cdot f_d(T_j, c_j) \cdot f_m(T_{jm}, c_{jm}) \cdot f_r(T_{jmr}, c_{jmr})$$

where

X_i	= vector of explanatory variables associated with zone i
f_g	= trip frequency choice model (Trip generation model)
f_d	= trip distribution choice model (Trip distribution model)
f_m	= mode choice model (Mode split model)
f_r	= route choice model (Traffic assignment model)

One way to think about the 4-stage model is in terms of what is called a general share formulation. Let:

T_j	= the number of trips produced from zone j
T_{ji}	= the number of trips from zone i to zone j
T_{jm}	= the number of trips from zone j to zone m , using mode m
T_{jmr}	= the number of trips from zone j to zone m , using mode m and route r (which may depend on j and m)
c_i	= denotes the cost of travel from zone i , taking account of all available modes and routes
c_{ji}	= denotes the cost of travel from zone j to each other zone, taking account of available modes and routes
c_{jm}	= denotes the cost of travel between zone j and zone m by each mode
c_{jmr}	= denotes the cost of travel between zone j and zone m along each available route by mode m from j to m

Travel Demand Models *

The 4-Stage Model (Aggregate Approach)

Stage 1: Trip Generation (Frequency Choice Model)

Stage 2: Trip Distribution (Destination Choice Model)

Stage 3: Mode Split (Mode Choice Model)

Stage 4: Traffic Assignment (Route Choice Model)

Objective: Estimate number of trips produced (P_j) and attracted (A_j) by each zone

Method: Regression Models. (Establish a statistical eqn between trip ends (P_j and A_j) and zonal attributes (eg. income, car ownership))

Formula: $P_j = f(X_i, c_i, \theta) + \epsilon$

'To' Zone			
'From' Zone	1	2	3
1			
2			
3			

A_1 A_2 A_3
 $(A_i$ means total trip attracted to zone i)
 $(P_j$ means total trip produced from zone j)

P_1
 P_2
 P_3
 we find all these $A_1, A_2, A_3, P_1, P_2, P_3$ in this stage.

Objective: Determining the trip interchanges between zones, completing the Trip Matrix, T_{ij} .

Methods: Gravity Models. Interaction is proportional to activity (P_i, A_j) and inversely proportional to separation/cost (:e. deterrence function)

doubly constrained model

Formula: $T_{ij} = P_i r_i A_j s_j f(c_{ij})$

(r_i and s_j are balancing factors solved iteratively to enforce row and column constraints.)

$f(c_{ij})$ is deterrence function, and typically is negative exponential:
 $f(c_{ij}) = e^{-\beta c_{ij}}$)

Objective: Split the Trip Matrix by mode (eg. car vs. bus). Find $T_{ijm} = T_{ij} \cdot Q_{ijm}$

Method: Logit Model. Find the proportion or probability of travelers using mode m , depends on the generalised cost of that mode relative to others.

Formula: $Q_{ijm} = \frac{T_{ijm}}{c_{ik} T_{ik}}$

$= \frac{f_m(c_{ijm})}{\sum_k f_k(c_{ijk})} = \frac{e^{-\beta c_{ijm}}}{\sum_k e^{-\beta c_{ijk}}}$

(β controls how sensitive mode choice is to cost)
 Note that " $-c_{ijm}$ " in this case is our utility function. If given V_s ... as our utility function:

$Q_{ijm} = \frac{e^{V_m}}{\sum_k e^{V_k}}$

Objective: Allocate Origin-Destination (OD) demand to specific routes/links on the network.

(unlike stage 1-3, this stage must account for congestion, where link costs rise as flow increases)

Methods:

1. All-or-Nothing (AON) assign all traffic to the minimum cost route. (Ignore congestion)

2. Wardrop's First Principle (User Equilibrium)

Traffic arranges itself so that all used routes have equal and minimum generalised (private) cost. (no individual driver can reduce the cost by switching lane)

3. Wardrop's Second Principle (System Optimal)

Traffic arranges itself to minimise the Total System Cost. (Require pricing / intervention to achieve: check e.g: congestion and cost part)

Activity-based Models (ABM) (Disaggregate Models)

Introduction

Core concept: Travel is derived from the need to participate in activities. The model schedules activities for individual rather than just counting aggregate trips.

Tours vs Trips: ABMs model Tours (chain of trips starting and ending at home) → maintaining consistency in time, space and mode.

Details: Uses high spatial (x-y coordinates) and temporal (1-minute resolution) disaggregation.

Why use ABM? *

1. Policy sensitivity: Better for analysing "soft" measures (eg. telecommuting, flex-time) and complex trade-offs.
2. Environmental precision: Can model "Cold start" vs. "Warm start" emissions based on how long a car has been parked between activities.
3. Interactions: Accounts for intra-household interactions (eg. who drops the kids at school?)

Comparison

Feature	4-Stage Model	Activity-Based Model (ABM)
Unit of Analysis	Aggregate Zones & Trips	Individual Agents & Activities/Tours
Consistency	Trips are independent (return trip not linked to outbound)	Ensures consistency (mode, time, location)
Data Needs	Zonal aggregate data	Aggregate individual data
Policy Use	Infrastructure & Pricing	Measures, complex behaviour, detailed missions
Weakness	Behaviourally simplistic; static	Complex to build; data intensive

Stage 1: Trip Generation Examples

Q1:

- The following table presents data collected in the last household O-D survey (made ten years ago) for three particular zones.

Zone	Residents/HH	Workers/HH	Mean income	Population
1	3	0.9	35000	30000
2	2.5	2.2	85000	50000
3	1.6	1.2	100000	10000

Ten years ago two household-based trip generation models were estimated using these data. The first is a linear regression model given by:

$$y = 0.1 + 0.6x_1 + 1.2Z_1$$

where y is the mean number of trips per household per day, x_1 is the number of workers in the household and Z_1 is a dummy variable which takes the value of 1 for high income ($>90\ 000$) households and 0 in other cases.

The linear regression model of trip generation estimated using household (hh) O-D survey data is as follows:

$$y = 0.1 + 0.6x_1 + 1.2Z_1$$

where y is the mean number of trips per household per day, x_1 is the number of workers in the household and Z_1 is a dummy variable which takes the value of 1 for high income ($>90\ 000$) households and 0 in other cases.

Zone 1 (average workers/hh=0.9, mean hh income=35000)
 $y = 0.1 + 0.6 * 0.9 + 1.2 * 0 = \mathbf{0.64 \text{ trips per household per day}}$

Zone 2 (average workers/hh =2.2, mean hh income =85000)
 $y = 0.1 + 0.6 * 2.2 + 1.2 * 0 = \mathbf{1.42 \text{ trips per household per day}}$

Zone 3 (number of workers=1.3, income=110000)
 $y = 0.1 + 0.6 * 1.3 + 1.2 * 1 = \mathbf{2.02 \text{ trips per household per day}}$

Therefore the number of trips generated in each zone can be calculated as
Number of trips per day = Trip rate y * Number of households
Number of trips per annum = Number of trips per day * 365

The results of these computations are presented in Table 3.

Table 3: Number of trips generated in each zone (calculated)

Zone	Number of Households	Trips per hh per day (Y)	Trips per day	Trips per annum	
1	10000	0.64	6400	2336000	<i>These are P₁, P₂ and P₃ respectively.</i>
2	20000	1.42	28400	10366000	
3	6250	2.02	12625	4608125	

(should check 2019 q3b)

Stage 2: Trip Distribution Examples

Q1:

Suppose we have a 3-zone system, in which the generalised cost of travel c_{ij} between zone i and zone j is given by the following matrix (in some suitable units, say pounds sterling):

'From' Zone	'To' Zone		
	1	2	3
1	0.2	1.2	2.4
2	3.8	0.2	5.7
3	0.5	3.2	0.2

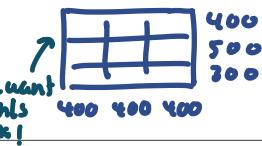
} this is for
 $f(c_{ij})$

For some future time period, a trip generation model predicts that P_i trips will be produced by zone i and A_j trips attracted to zone j . The values of P_i and A_j are as follows:

$$\begin{array}{ll} P_1 = 400 & A_1 = 400 \\ P_2 = 500 & A_2 = 400 \\ P_3 = 300 & A_3 = 400 \end{array}$$

The deterrence function is of the form $f(c_{ij}) = e^{-\beta c_{ij}}$, where $\beta = 0.02$.

what we know for now is
only the sum of row and column :



(a)

- Compute the trip matrix generated by the 'Unconstrained' gravity model:

$$T_{ij} = K P_i A_j f(c_{ij})$$

where K is a global scaling constant such that $\sum_i T_{ij} = 1200$

Task 1 Unconstrained gravity model

The idea of the gravity model is that the propensity for interaction between two zones is proportional to the product of a measure of the activity in each zone divided by a measure of the separation (or travel cost) between the zones. That is:

$$T_{ij} \propto P_i A_j f(c_{ij})$$

and therefore

$$T_{ij} = K P_i A_j f(c_{ij})$$

where K is a global scaling constant such that: $\sum_i T_{ij} = 1200$

We first compute $T'_{ij} = P_i A_j f(c_{ij})$ as follows:

$T_{11}' = P_1 A_1 f(c_{11}) = 400 * 400 * \exp(-0.02*0.2) = 159361.278$
and so on.

	1	2	3
1	159361.278	156205.714	152501.406
2	185363.241	199201.598	178451.591
3	118805.980	112560.600	119520.959
	463530.500	467967.911	450473.95

first step is always to fill the trip matrix based on A_j , P_i and $f(c_{ij})$, doesn't matter what constraint model we use.
(i.e. ignore s_j , r_i for now)

K is therefore calculated as $\sum_i T_{ij} / \sum_i T_{ij}' = 1200 / 1381972.367 = 0.003222$

we just scale the sum sum!

And the forecast trip matrix, $T_{ij} = K T'_{ij}$, which is calculated as follows (rounded to 0 decimal places):

T_{ij}	1	2	3	
1	138	136	132	406
2	161	173	155	489
3	103	98	104	305
	402	406	391	1200

what we do in unconstrained gravity model is make this total sum = T_{ij}
 $= A_1 + A_2 + A_3 = P_1 + P_2 + P_3 = 1200$

Note that using this method the row and column sums do not satisfy the production and attraction targets (row- and column-sums), but the total number of trips generated and attracted, i.e. $\sum_i T_{ij}$, meets the target of 1200. This is the property of the unconstrained gravity model, as it is constructed to only meet the total trips constraint.

(b)

- 2 Compute the trip matrix generated by the 'Singly constrained' gravity model:

$$T_{ij} = P_i r_i A_j f(c_{ij})$$

where the r_i are constants such that $\sum_j T_{ij} = P_i \quad \forall i$

Task 2 Singly constrained gravity model

We are asked to apply the singly constrained gravity model with the constraint on the number of trip destinations (or attractions). That is:
 $\sum_j T_{ij} = P_i$

This constraint leads to the following model:
 $T_{ij} = P_i r_i A_j f(c_{ij})$

where r_i are constants that satisfy the constraint $\sum_j T_{ij} = P_i$ for all j and s_j is given by:
 $r_i = [\sum_j A_j f(c_{ij})]^{-1}$

So we first calculate $T'_{ij} = P_i A_j f(c_{ij})$ as given below, i.e.
 $T'_{11} = P_1 A_1 f(c_{11}) = 400 * 400 * \exp(-0.02*0.2) = 159361.278$
and so on.

	1	2	3	
1	159361.278	156205.714	152501.405	468068.398
2	185363.241	199201.598	178451.591	563016.430
3	118805.980	112560.600	119520.959	350887.539
	463530.500	467967.911	450473.956	1381972.367

Note that this is the same as the first step in the unconstrained gravity model calculation.

The next step is to scale the elements in the T'_{ij} matrix by r_i , where
 $r_1 = 1 / [A_1 * f(c_{11}) + A_2 * f(c_{12}) + A_3 * f(c_{13})] = 0.000855$
 $r_2 = 1 / [A_1 * f(c_{21}) + A_2 * f(c_{22}) + A_3 * f(c_{23})] = 0.000888$
 $r_3 = 1 / [A_1 * f(c_{31}) + A_2 * f(c_{32}) + A_3 * f(c_{33})] = 0.000855$

and $T_{11} = T'_{11} * r_1$ and so on.

The forecast trip matrix is therefore as follows

	1	2	3	
T_{ij}	136.186	133.490	130.324	400.000
	164.616	176.906	158.478	500.000
	101.576	96.236	102.187	300.000
	402.379	406.632	390.990	1200.000

The final forecast trip matrix rounded up to 0 decimal places is as follows.

	1	2	3	
T_{ij}	136	133	130	400
	165	177	158	500
	102	96	102	300
	402	407	391	1200

Note that using this method we satisfy the row constraints i.e. the number of trips generated from a zone, by design. But we do not meet the column targets of trip attractions.

first step again is fill the trip matrix.

in singly constrained model,
Instead scaling the TOTAL sum,
we scale the sum of row OR sum
of column, based on the question.

this question wants:

$T_{ij} = P_i r_i A_j f(c_{ij})$
we scale P_i (sum of row)

(c)

- 3 Compute the trip matrix generated by the 'Doubly constrained' gravity model:

$$T_{ij} = P_i r_i A_j s_j f(c_{ij})$$

where the r_i are constants such that $\sum_j T_{ij} = P_i \forall i$ and s_j are constants such that $\sum_i T_{ij} = A_j \forall j$

Task 3 Doubly constrained gravity model

The doubly constrained gravity model is designed to meet both the production and attraction constraints i.e. $\sum_i T_{ij} = A_j$ and $\sum_j T_{ij} = P_i$

This leads to the model:

$$T_{ij} = P_i r_i A_j s_j f(c_{ij})$$

where

r_i constants that satisfy the constraint $\sum_j T_{ij} = P_i$ for all i

s_j constants that satisfy the constraint $\sum_i T_{ij} = A_j$ for all j

These are given by:

$$r_i = [\sum_j A_j s_j f(c_{ij})]^{-1}$$

$$s_j = [\sum_i P_i r_i f(c_{ij})]^{-1}$$

Since r_i is a function of s_j and vice versa, this does not have a direct solution and can be solved only through an iterative method.

ITERATION 1

In iteration 1 we are required to meet the production constraints i.e. $\sum_j T_{ij} = P_i$

So we start with the convenient assumption that $s_j=1$ for all j , which yields the following model

$$T_{ij} = P_i r_i A_j f(c_{ij})$$

where

$$r_i = [\sum_j A_j f(c_{ij})]^{-1}$$

This is the same as the singly unconstrained model that meets attraction constraints. Therefore at the end of iteration 1 we have the trip matrix from part (b)

$T_{ij}(1)$	1	2	3	
<i>Iteration 2:</i>	136.186	133.490	130.324	400.000
2	164.616	176.906	158.478	500.000
3	101.576	96.236	102.187	300.000

scale column by $\frac{400}{402.379}$ scale column by $\frac{400}{406.632}$ scale column by $\frac{400}{390.990}$

this is after scaling row.

In iteration 2, we use the output of iteration 1, i.e. $T_{ij}(1)$, and calculate $T_{ij}(2)$ that meets the attraction constraints i.e. $\sum_i T_{ij} = A_j$. That is:

$$T_{ij}(2) = A_j * [T_{ij}(1) / \sum_k T_{kj}(1)]$$

$$= 400 * [136.186 / (136.186 + 164.616 + 101.576)] = 135.381$$

and so on.

The result of iteration 2 is as follows

	1	2	3	
$T_{ij}(2)$	135.381	131.313	133.327	400.021
	163.643	174.021	162.130	499.794
	100.976	94.667	104.542	300.185
	400.000	400.000	400.000	1200.000

ITERATION 3

In iteration 3, we use the output of iteration 2, i.e. $T_{ij}(2)$, and calculate $T_{ij}(3)$ that again meets the production constraints i.e. $\sum_j T_{ij} = P_i$. That is:
 $T_{ij}(3) = P_i * [T_{ij}(2)/\sum_k T_{ik}(2)]$, which computes as

$$T_{11}(3) = P_1 * [T_{11}(2)/[T_{11}(2)+T_{12}(2)+T_{13}(2)]] \\ = 400 * [135.381/(135.381+131.313+133.327)] = 135.374$$

and so on.

The result of iteration 3 is as follows:

	1	2	3	
$T_{ij}(3)$	1	135.374	131.306	133.320
	2	163.711	174.092	162.197
	3	100.913	94.609	104.478
		399.998	400.006	399.995
				1200.000

ITERATION 4

Finally, in iteration 4, we use the output of iteration 3, i.e. $T_{ij}(3)$, and calculate $T_{ij}(4)$ that meets the attraction constraints i.e. $\sum_i T_{ij} = A_j$. That is:
 $T_{ij}(4) = A_j * [T_{ij}(3)/\sum_k T_{ij}(3)]$ which computes as:

$$T_{11}(4) = A_1 * [T_{11}(3)/[T_{11}(3)+T_{21}(3)+T_{31}(3)]] \\ = 400 * [135.374/(135.374+163.711+100.913)] = 135.375$$

and so on.

The result of iteration 4 is as follows

	1	2	3	
$T_{ij}(4)$	1	135.375	131.303	133.322
	2	163.711	174.089	162.199
	3	100.914	94.607	104.479
		400.000	400.000	400.000
				1200.000

The final trip matrix using the doubly constrained model, and rounded up to 0 decimal places is

	1	2	3	
T_{ij} -final	1	135	131	133
	2	164	174	162
	3	101	95	104
		400	400	400
				1200

During each of the iterations the trip matrix satisfies one of either the row or column constraints exactly. And as seen from the outputs above, with each iteration the results come closer to convergence i.e. to satisfying both the row and column constraints simultaneously. At the end of iteration 4, we almost satisfy both constrain simultaneously while at the same time taking into consideration the cost of travel between each zone.

Stage 3 : Mode Split Example

Q3.

A binomial logit model has been calibrated and the following utility functions obtained:

$$V_c = 0.4 - 0.40t_i - 0.90t_o - 0.032c_c - 0.08c_p + 1.2M + 0.05I$$

$$V_r = -0.35t_i - 0.95t_o - 0.08c_r$$

where V_c is the utility function for the car mode

V_r is the utility function for the rail mode

t_i is the in-vehicle time in minutes

t_o is the out-of-vehicle time in minutes

c_c is the cost of travel by car in pence

c_p is the cost of car parking in pence

c_r is the cost of travel by rail in pence

M is a dummy mode that takes a value of 0 for males and 1 for females

I is the income in £1000s per year

1. Interpret the value of the parameters in the utility functions above.

2. For a particular man with an income of £50,000 per year, the variables have the following values:

	Car	Rail
In-vehicle time in minutes	15	12
Out-of-vehicle time in minutes	5	10
Cost of travel in pence	80	120
Cost of parking in pence	150	-

(a) What would be the probability of him travelling by car?

(b) What would be the probability of a woman with the same salary travelling by car?

(c) What would the probabilities of travelling by car for each of them change to if the cost of parking was increased to 180 pence?

2.

(a) For a particular man (i.e. $M = 1$) with

Annual income = $I = £50,000 = 50$ (in £1000s)

In-vehicle travel time; t_i (car) = 15 min, t_i (rail) = 12 min

Out-of-vehicle travel time; t_o (car) = 5 min, t_o (rail) = 10 min

Cost of travel by car = $c_c = 80$ pence

Cost of travel by rail = $c_r = 120$ pence

Cost of parking = $c_p = 150$ pence

Therefore:

$$\begin{aligned} V_{car} &= 0.4 - 0.40t_i - 0.90t_o - 0.032c_c - 0.08c_p + 1.2M + 0.05I \\ &= 0.4 - 0.40 \cdot 15 - 0.90 \cdot 5 - 0.032 \cdot 80 - 0.08 \cdot 150 + 1.2 \cdot 1 + 0.05 \cdot 50 \\ &= -20.960 \end{aligned}$$

$$\begin{aligned} V_{rail} &= -0.35t_i - 0.95t_o - 0.08c_r \\ &= -0.35 \cdot 12 - 0.95 \cdot 10 - 0.08 \cdot 120 \\ &= -23.300 \end{aligned}$$

Probability of the man travelling by car = $\exp(V_{car}) / [\exp(V_{car}) + \exp(V_{rail})] = 0.912$

(b) The probability of a woman in the same circumstances travelling by car is calculated as follows:

All the variables are the same except $M=0$

$$V_{car} = -22.160$$

$$V_{rail} = -23.300$$

Probability of the woman travelling by car = $\exp(V_{car}) / [\exp(V_{car}) + \exp(V_{rail})] = 0.758$

In other words, all else being equal, this model suggests that a woman is less likely to travel by car than a man. This is due to the negative gender (male) variable in the disutility for rail, suggesting that men inherently assign a higher disutility to travelling by rail than women.

(c) Now, if the cost of parking was increased to 180 pence i.e. $c_p = 180$ pence

For the man (i.e. $M = 1$)

$$\begin{aligned} V_{car} &= 0.4 - 0.40t_i - 0.90t_o - 0.032c_c - 0.08c_p + 1.2M + 0.05I \\ &= 0.4 - 0.40 \cdot 15 - 0.90 \cdot 5 - 0.032 \cdot 80 - 0.08 \cdot 180 + 1.2 \cdot 1 + 0.05 \cdot 50 \\ &= -23.360 \end{aligned}$$

$$\begin{aligned} V_{rail} &= -0.35t_i - 0.95t_o - 0.08c_r \\ &= -0.35 \cdot 12 - 0.95 \cdot 10 - 0.08 \cdot 120 \\ &= -23.300 \end{aligned}$$

Probability of the man travelling by car = $\exp(V_{car})/[\exp(V_{car}) + \exp(V_{rail})] = 0.485$

For the woman (i.e. $M = 0$)

$$\begin{aligned} V_{car} &= 0.4 - 0.40t_i - 0.90t_o - 0.032c_c - 0.08c_p + 1.2M + 0.05I \\ &= 0.4 - 0.4*15 - 0.90*5 - 0.032*80 - 0.08*180 + 1.2*0 + 0.05*50 \\ &= -24.560 \end{aligned}$$

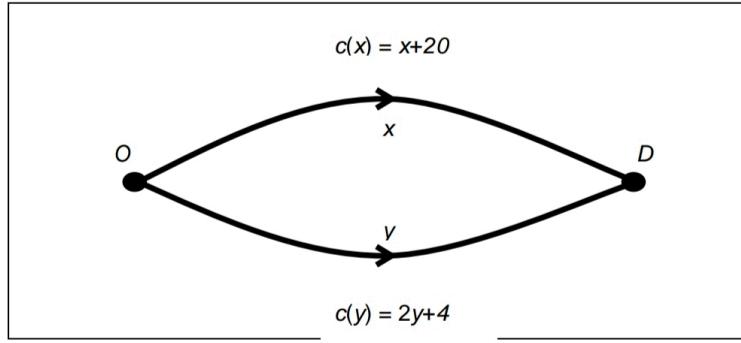
$$\begin{aligned} V_{rail} &= -0.35t_i - 0.95t_o - 0.08c_r \\ &= -0.35*12 - 0.95*10 - 0.08*120 \\ &= -23.300 \end{aligned}$$

Probability of the man travelling by car = $\exp(V_{car})/[\exp(V_{car}) + \exp(V_{rail})] = 0.221$

Stage 4 : Traffic Assignment Model

Q1.

Suppose we have a system comprising a single origin O connected to a single destination D via a congestable network consisting of two disjoint routes, each represented by a single link, as shown below.



The demand for travel between O and D is q and this demand induces flows of x on the upper route and y on the lower route.

The generalised (private) cost of using the upper route is $c(x) = x+20$ and generalised (private) cost of using the lower route is $c(y) = 2y+4$.

The demand q between O and D is fixed at 80.

TASKS

- 1 Compute the traffic assignment corresponding to a Wardropian (WP1) equilibrium in this network and hence find the equilibrium travel time between O and D and the total system cost.
- 2 In what ways would you expect a system optimising (WP2) flow pattern in this network to differ from the WP equilibrium?

SOLUTION

Task 1 (WP1's example) — $c(x)=c(y)$

The solution to the Wardropian equilibrium problem (WP1) is that assignment which results in equal journey time along each route.

Therefore at a WP1 equilibrium

$$x+20 = 2y+4 \quad (1)$$

and by conservation of flow we also have

$$x+y=80 \quad (2)$$

Equations (1) and (2) are a pair of linear simultaneous equations in two variables, and are easily solved by direct substitution (e.g. for y in (1) from (2)). This gives,

$$\begin{aligned} x+20 &= 2(80-x)+4 \\ \Rightarrow x &= 48, \text{ and } y = 32. \end{aligned}$$

So in the WP1 solution, the cost along each route is 68 and the total network cost TC is $48*68 + 32*68 = 80*68$. Hence $TC_{WP1} = 5440$.

Task 2 (WP2's example) — $T = xc(x) + yc(y)$ $\frac{\partial T}{\partial x} = 0$ or $\frac{\partial T}{\partial y} = 0$

An assignment according to Wardrop's Second Principle is one in which the total network cost is minimised.

For a given assignment, the total network cost TC is

$$TC = x(x+20) + y(2y+4) \quad (3)$$

Using the conservation equation (2), we can express TC in terms of either x or y alone.
Expressed in terms of x ,

Hence

$$TC(x) = x(x+20) + (80-x)[2(80-x)+4]$$

$$TC(x) = 3x^2 - 304x + 13120 \quad (4)$$

TC is minimised when $\frac{\partial TC(x)}{\partial x} = 0$ and $\frac{\partial^2 TC(x)}{\partial x^2} > 0$

$$\frac{\partial TC(x)}{\partial x} = 6x - 304 \text{ and } \frac{\partial^2 TC(x)}{\partial x^2} = 6 (> 0)$$

Hence $x = 50.7$ and $y = 29.3$ is a minimum.

Notice that in the WP2 solution the costs along each route are not equal. Along the upper route the cost is 70.7 and along the lower route it is 62.6. Hence, relative to the WP1 equilibrium, travellers on the upper route are worse off but travellers along the lower route are better off.

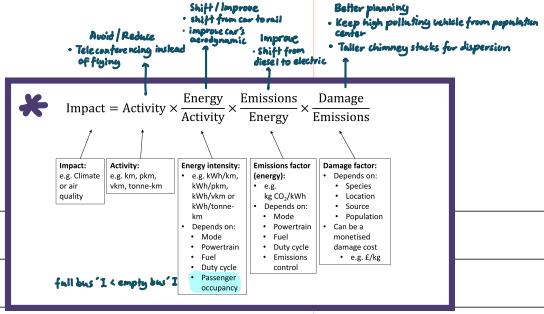
Substituting into (4) gives $TC_{WP2} = 5418.7$.

As we would expect, $TC_{WP1} > TC_{WP2}$.

Hence, despite the fact that in a WP1 solution each individual seeks to achieve the best possible outcome for themselves, the aggregate outcome for everybody is not optimal. The reason for this difference is the divergence between individual (private) costs and marginal (social) costs. Intuitively, what happens is that in the attempt to achieve the best possible outcome for themselves each driver interferes with the passage of other drivers, and does so more than is absolutely necessary. When all drivers behave in this way the system is in aggregate worse off.

In this case in the WP1 solution, too many drivers ($=50.7-48=2.7$) choose to use the lower route and although these particular individuals are better off using this route than the upper route, they cause extra congestion for themselves and all the other users on the lower route.

c9. Transport Environment



Transport Environment

Transport Energy

1. Key Physics and Equations

$$F_{\text{drag}} = \frac{1}{2} \rho A_f C_d V^2 \quad (\text{Aerodynamic drag})$$

air density frontal drag area speed relative to air

$$F_{\text{roll}} = CRmg \quad (\text{Rolling resistance})$$

coeff of rolling resistance gravity
 (depends on tire/road only!) mass of vehicle

$$F_{\text{grad}} = mgs \sin \theta \quad (\text{Gradient Force})$$

Weight of the vehicle in the sloped direction.

2. Work and Power

Work (Energy) = Force × Distance

(unit: J or kWh)

Power (Rate of Energy Consumption)

$$= F_{\text{total}} \times V \quad (\text{velocity})$$

(note: $F \propto V^2$, hence the power required to overcome drag is proportional to V^3 , hence speed limit saves fuel)

3. Energy Intensity ($I = \underline{\text{AIED}}$)

used to compare energy efficiency across different modes (e.g. Rail vs Car)

$$\text{Energy Intensity} : \frac{\text{Energy consumed (kWh)}}{\text{Activity (unit: km, plkm, vkm)}}$$

Combustion and Emissions

1. Combustion

Ideal: $\text{Fuel (C}_x\text{H}_y\text{)} + \text{Air} \rightarrow \text{CO}_2 + \text{H}_2\text{O} + \text{Heat}$



Real: Usually incomplete combustion, which leads to by products such as:

- CO (carbon monoxide): incomplete oxidation
- UHC (unburned hydrocarbon)
- PM (particulate matter): soot/black carbon (incomplete oxidation) + non-exhaust sources (brake/tyre wear)
- NO_x (nitrogen oxide): formed at high temp. when N₂ in air oxidises (NO nitric oxide, NO₂ nitrogen dioxide)

2. Greenhouse Gases (GHGs) vs. Air Pollutants

GHGs (Global Impact): Causes climate change

- primary CO₂, CH₄, N₂O
- Global Warming Potential (GWP) is a metric to compare GHGs to CO₂

$$\text{CO}_2\text{e} = \sum (\text{Mass}_i \times \text{GWP}_i)$$

e.g. CO₂e of CO₂ = 1, CO₂e of CH₄ = 28 (over 100 yrs)

GWP depends on timeframe!

Pollution Dispersion

1. Simple Box Model

We assume that the street is a "box". We want to solve for steady state concentration (X)

$$\text{Formula: } X = \bar{X}_B + \frac{L \times q}{U \times H}$$

X : conc. in the street (µg/m³ or kg/m³)

\bar{X}_B : background conc.

L : length of the street.

H : mixing height (often taken as height of building)

U : wind speed parallel to the st.

q : emission rate per unit area (kg/m²s)

3. Emission Factors ($I = \underline{\text{AEID}}$)

- Can be energy based form or distance based form

$$\text{Impact} = \text{Activity} \times \frac{\text{Energy}}{\text{Activity}} \times \frac{\text{Emissions}}{\text{Energy}} \times \frac{\text{Damage}}{\text{Emissions}}$$

energy based form:

$$\text{emission factor} = \frac{\text{emission}}{\text{energy}} \quad (\text{unit: kgCO}_2/\text{kWh})$$

$$\text{Impact} = \text{Activity} \times \frac{\text{Energy}}{\text{Activity}} \times \frac{\text{Emissions}}{\text{Energy}} \times \frac{\text{Damage}}{\text{Emissions}}$$

distance based form:

$$\text{emission factor} = \frac{\text{emission}}{\text{activity}} \quad (\text{unit: kgCO}_2/\text{km})$$

Impact Pathway

Transport activity



Energy consumption



Emissions (e.g. kg of GHG or other pollutants)

} combustions

Atmospheric dispersion and processing



Impact and response (e.g. climate, human health)



Monetary evaluation of costs (£££)