# 物流系统分析
# Logistics System Analysis

## 第 6 周　一到一配送问题（2）–其他一维问题
## One-to-One Distribution–Other 1-D Logistics Problems

葛乾

西南交通大学　系统所 & 交通工程系

# 目 录 |Outline


1. Other One-Dimensional Location Problems
2. Accuracy of the CA Expression
3. Generalization of the CA Approach

# Bus departure schedule problem

The CA technique was originally proposed to find a near-optimal bus departure schedule from a depot (Newell, 1971).
Given the cumulative number of people $D(t)$ demanding service by time t, the fixed cost of a bus dispatch $c_f$, and the cost of each person-hour waited $c_i$, the objective was to minimize the sum of the bus dispatch (motion) and waiting (holding) costs.
With an unlimited bus capacity, this problem is almost identical to the one we have just solved; except for $D(t)$, which now represents the cumulative number of people (items) entering the system and not the number leaving. The cost equations still hold.

# Freight terminals location on a distance line

- This problem locates freight terminals on a distance line between 0 and $d_{\max}$. This interval contains origins, which send items to a depot. The distance line extends from the origin, O, to a depot, located at $d = \tilde{d} \geqslant d_{\max}$.
- The flow of freight (number of items per day) that originates between $O$ and $d$ is a function of $d$, $D(d)$, which increases from 0 to $v_{tot}$. Items are individually carried to the terminals at a cost $c'_d$ per unit distance per item. Each day a vehicle travels the route collecting the items accumulated at each terminal and takes them to the depot.

# The access cost

- The motion cost for this operation has three components: the handling cost at the terminals, assumed to be constant and therefore ignored, the access cost to the terminals, and the line-haul cost of operating the vehicle from the terminals to the depot. 构成：枢纽处的处理成本，可视为常数并忽略；到达枢纽的成本；干线运输的成本，即从枢纽直运到目的地的费用
- The access cost is given by the product of $c'_d$ and the total item-miles of access traveled per day; it increases with the separation between stops.

# The line-haul cost

The line-haul cost has the form of

$$\text{line-haul cost/day} = c_s(1 + n_s) + c_d(\tilde{d}) + c_s'(v_{\text{tot}})$$

- $c_s$, is the cost attributable to each trip, regardless of distance and shipment composition; it includes the cost of stopping the vehicle and having it sit idle while it is being loaded and unloaded. Think of it as the fixed cost of stopping, independent of what is being loaded and unloaded.
- $n_s$ is the number of stops (excluding the depot)
- $c_d$ is the cost of vehicle-mile. It is the vehicle cost (including the driver) for distance traveled regardless of the vehicle's contents;
- $c_s'$, represents the cost of carrying items. It represents a penalty for delaying the vehicle while loading and unloading the items, as well as the cost of handling the items within the vehicle. $v_{\text{tot}}$ is the total size of the shipment arriving at the depot.

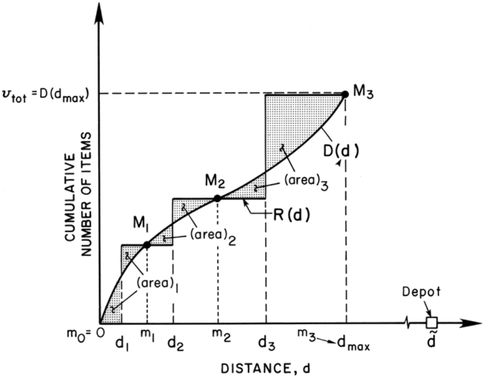
# The line-haul cost (cont.)

Note that the line-haul cost does not depend on the specific stop locations and that in contrast to the access cost, it increases with $n_s$.

$$\text{line-haul cost} = c^o + c_s n_s$$

where $c^o$ is a constant that will be ignored for design purposes.
As the problem has been formulated, with one trip per day, the sum of the holding costs at all stops can be ignored - consideration reveals that the sum is constant. Pipeline inventory costs do depend on the decision variables (they should increase with $n_s$) but for cheap freight the effect is negligible. Thus, all inventory and holding costs are neglected.
**The stops will be located as the result of a trade-off between line-haul and access costs.** Without this simplification, which is inappropriate for passenger transportation, the problem is equivalent to the transit stop location problem.

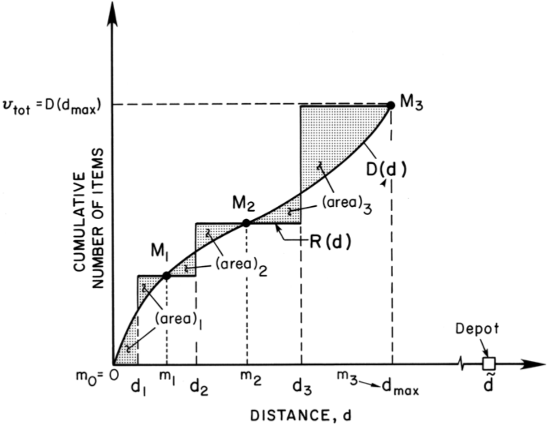
# 示意图




The figure depicts the location of three terminals (at points $d_1$, $d_2$, and $d_3$) and a curve, $R(d)$, depicting the number of items in the vehicle as a function of its position. This curve increases in steps at each terminal location. The size of each step equals the number of items collected.

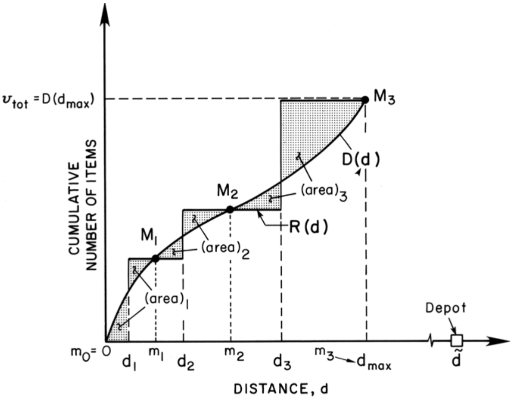
To minimize access (and total) cost each item is routed to the nearest terminal, and as a result the step curve passes through the midpoints, $M_i$, shown in the figure.

思考一下，为什么需要考虑中转站的中点?




计算中转站的中点是为了划分每个枢纽的服务范围. 例如, $[m_0, m_1]$ 范围内的货物由在 $d_1$ 处中转, $[m_1, m_2]$ 范围内的货物由在 $d_2$ 处中转等, 因为对应的枢纽距离货物位置最近。

The coordinates of $M_i$ are $m_i = (d_i + d_{i+1})/2$ and $D(m_i)$; with $m_0 = 0$ and $m_{n_s} = d_{\max}$




Let us see how the total cost can be prorated to short intervals, by considering the partition of $(0, d_{\max}]$ into the following intervals surrounding each terminal: $I_1 = (0, m_1], I_2 = (m_1, m_2], \ldots, I_{n_s} = (m_{n_{s-1}}, d_{\max}]$. Each interval, $I_i$, adds an access cost proportional to the daily item-miles traveled for access to terminal $i$.

This is given by the shaded area on the two quasi-triangular segments next to the location of the terminal, $(area)_i$, thus

$$\text{access cost}_i = (area)_i c'_d. \qquad 平均的运送距离是(m_i - m_{i-1})/4$$

For slowly varying $D(d)$, the access cost can be rewritten as:

$$\text{access cost}_i \approx \frac{1}{4}(m_i - m_{i-1})^2 D'(d_i) c'_d.$$

Since each terminal adds $c_s$ to the daily line-haul cost, the share of the total cost prorated to $I_i$ is:

$$(\text{Total cost per day})_i \approx c_s + \frac{c'_d}{4}(m_i - m_{i-1})^2 D'(d_i).$$

Since $D'(d) \approx D'(d_i)$ for $d \in I_i$ (we stated that $D'(d)$ varied slowly), the above expression can be approximated by:

$$(\text{Total cost per day})_i \approx \int_{m_{i-1}}^{m_i} \{\frac{c_s}{m_i - m_{i-1}} + \frac{c'_d}{4}(m_i - m_{i-1}) D'(d)\} \mathrm{d}d.$$

If we now let $s(d)$ denote a slowly varying function such that $s(d_i) = m_i - m_{i-1}$ (the function, used later to locate the terminals, indicates the size of a terminal's influence area depending on location), then we can rewrite the last expression once again, using $s(d)$ instead of $mi - m_{i-1}$:

$$(\text{Total cost per day})_i \approx \int_{m_{i-1}}^{m_i} \{\frac{c_s}{s(d_i)} + \frac{c_d'}{4}(s(d_i))D'(d)\}\mathrm{d}d.$$

The total cost for the system is then:

$$(\text{Total cost per day}) \approx \int_{0}^{d_{\max}} \{\frac{c_s}{s(d)} + \frac{c_d'}{4}(s(d))D'(d)\}\mathrm{d}d.$$

# Analytical form of EOQ

The least cost $s(d)$ minimizes the integrand at every point; given its EOQ analytical form, we find

$$s(d) \approx 2[\frac{c_s}{c'_d D'(d)}]^{1/2}.$$

Note that if $D'$ varies slowly, $s(d)$ will vary slowly as we had assumed.
The expressions for the minimum total and average (per item) cost as follows

$$\text{(Total cost per unit time)}^* \approx \int_0^{d_{\max}} [c_s c'_d D'(d)]^{1/2} \mathrm{d}d$$

$$\text{cost per item}^* \approx \int_0^{d_{\max}} [c_s c'_d D'(d)]^{1/2} D'(d) \mathrm{d}d \Big/ \int_0^{d_{\max}} D'(d) \mathrm{d}d$$

# Locate the terminals

- To locate the terminals, one first divides $(0, d_{\max}]$ into non-overlapping intervals of approximately correct, length $I_1, I_2$, etc.. . . , by starting at one end and using the EOQ formula repeatedly.
- If the last interval is not of correct length, then the difference can be absorbed by small changes to the other intervals.
- If $d_{\max}$ is large (so that there are at least several intervals), then the final partition should satisfy $s(d) \approx m_i - m_{i-1}$ if $d \in I_i$, and the approximations leading to the equations should be valid.
- With the influence areas defined in this manner, the terminals are located next. They should be positioned within each interval so that the boundary between neighboring intervals is equidistant from the terminals.
- For a general sequence of intervals (e.g., of rapidly fluctuating lengths) this may be difficult (even impossible) to do, but for our problem with $|I_i| \approx |I_{i+1}|$ the best locations should be near the center of each interval; in fact little is lost by locating the terminals at the centers.

# 目 录 |Outline

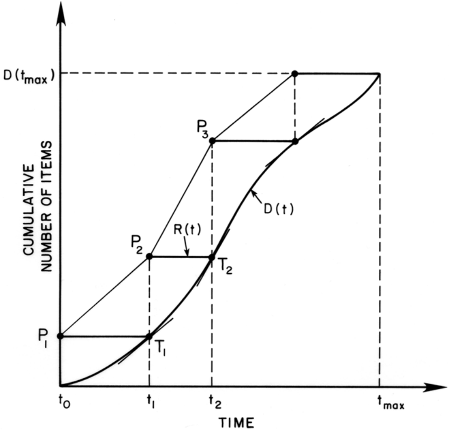

1. Other One-Dimensional Location Problems

2. Accuracy of the CA Expression

3. Generalization of the CA Approach

- Although a systematic analysis of its errors has not been reported, experience indicates that the CA approach is very accurate when the descriptive characteristics of the problem ($D'(t)$ in the text's examples) vary slowly as assumed.
- The approach is also robust. It is effective even if the variation in conditions is fairly rapid –in our case, accurate results are obtained even if $D'(t)$ varies by a factor of two within the influence areas. This conclusion is not surprising in light of the EOQ robustness discussed in the previous lecture
- When conditions are unfavorable, the CA method can both over- and under-predict the optimal cost. The textbook provides two examples identify said conditions, with the first example illustrating overestimation and the second underestimation. The basis for comparison will be the exact solution, which for our problem can be obtained readily.

# Recall the construction method




Recall our previous construction method for the cumulative number of items shipped versus time

- Choose a point $P_1$ on the ordinates axis and move across to $T_1$
- Draw from $P_1$ a line parallel to the tangent to $D(t)$ at $T_1$, and draw from $T_1$ a vertical line. Label the point of intersection $P_2$ .

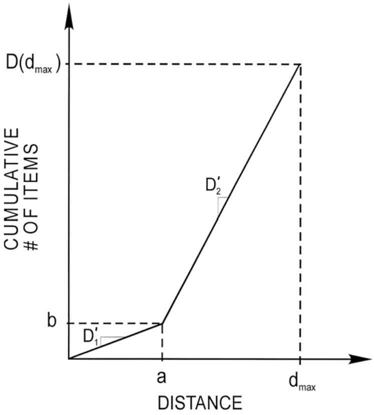
之前例子中横轴是时间，中转枢纽问题中横轴是距离

# Construction method for the terminal location problem

- Given # of stops $n_s$, for a set of locations to be optimal the line $D(d)$ of the distance figure must bisect in two equal halves every vertical segment of $R(d)$. Otherwise, the terminal (e.g., terminal 3) could be moved slightly to decrease access cost. The optimal solution can then be found by comparing all the possible $R(d)$ with the above property.
- For a given $d_1$, draw a vertical step that is bisected by $D(d)$, and move across horizontally so that the horizontal segment is also bisected by $D(d)$. This identifies $d_2$. Repeat the construction to find $d_3$, $d_4$, etc. (Only those values of $d_1$ for which the last vertical segment is bisected by $D(d)$ need to be considered seriously.) The optimal solution corresponds to a $d_1$ which minimizes the sum of the stop cost and access cost.
- The procedure is so simple that it can be implemented in spreadsheet form*.

*The user selects $d_1$ and the spreadsheet returns the graphs, and the cost; it is then easy to find the solution either interactively or automatically with the computer.

# Example 1: Overestimation




Terminals are to be located on two adjoining regions with high and low demand. The left figure depicts a generic piece-wise linear cumulative demand curve of this type. The coordinates of the break-point (distance, item number) are given by parameters "a" and "b". They must be consistent with the specified values for $d_{\max}$, $D(d_{\max})$, $D'_1$ and $D'_2$. For this problem the continuum approximation approach yields

$$TC^* \approx (c_s c'_d)^{1/2}\{a\sqrt{D'_1} + (d_{\max} - a)\sqrt{D'_2}\}$$

# Parameters

A possible set of parameters is $d_{\max} = 500, D(d_{\max}) = 1700, D_1' = 1, D_2' = 5, a = b = 200, c_d' = 1$ and $c_s = 160,000$. This choice has been made because a systematic analysis shows that it produces the largest overprediction error in percentage terms. The predicted cost is: $TC^* = 348,328$.
In actuality the least possible cost is 8% smaller. It arises when a single terminal is located at $d = 330$. The reader can verify that the exact access cost for this location is 160,500 units. Since the terminal cost is 160,000 units (for one terminal), the grand total is 320,500 $<$ 348,328.

# Insights from the example

This rather extreme example illustrates that the CA approach can overestimate the optimum cost.

- To understand why this happens let us decompose the CA costs into its components. Note first that the ideal spacing between terminals predicted by the CA method with is:
  $s(d) = 2 * [160,000/(1*1)]^{1/2} = 800$ in the low demand section, and
  $s(d) = 2 * [160,000/(1*5)]^{1/2} \approx 357$ in the high demand section.
- The CA access cost is calculated as if the average access distance was $s(d)/4 = 200$ in the low demand section and 89.25 units in the high demand section. Since there are 200 items in the low density region and 1500 in the high density region, the total CA access cost is approximately: $200 \times 200 + 89.25 \times 1500 = 173,875$.
- The CA stop cost (line-haul cost) is calculated by integrating the density of terminals over the service region, $(200/800 + 300/357) \approx 1.09$, and multiplying this result by the cost of a terminal: $1.09 \times 160,000 = 174,400$.
- The grand total is therefore: $173,875 + 174,400 = 348,275 \approx 348,328$.

# Insights from the example (cont.)

It turns out, however, that just a single terminal in the high density region can serve both, the low density points with an average distance barely greater than the CA access distance, and the high demand section with an average access distance considerably inferior to the corresponding CA distance. For our chosen location ($d = 330$) the actual average access distances are: 230 units for the low density section (200 with the CA method)and 76 for the high density section (89 with CA method). Since we are using only one terminal, the final cost is lower.

The overprediction effect arises because the demand curve varies significantly and very favorably between the terminal and the edge of the service region, and the CA approach does not exploit this variation. The variation is so favorable that it allows a terminal provided for the high density points to double up efficiently as a terminal for the low density points.

Favorable conditions are unusual, however. When the demand does not vary rapidly the CA approach consistently underestimates demand.

# Example 2: Underestimation example

- By its nature, the CA approach ignores that the number of terminals must be an **integer**; any situation with a finite region size (or time horizon) will exhibit this error type.
- To exclude the overprediction error type illustrated by example 1, the demand per unit length of region is set constant: $D'(d) = D'$. This also allows closed form comparisons to be made.
- The CA solution is $TC^* = \sqrt{c_s c_d'}\sqrt{D'} d_{\max}$. Without losing generality, we choose the units of distance, item quantity and money so that $d_{\max} = 1, D(d_{\max}) = 1$ and $c_s = 1$. Thus, $D' = 1$ and only the parameter $c_d'$ remains. The above expression becomes: $TC^* = \sqrt{c_d'}$
- If the exact optimal solution has $n_s$ terminals, the distance line will be partitioned into $n_s$ intervals of equal length: $I_i = ((i-1)/n_s, i/n_s]$. The total cost is then

$$TC(n_s) = n_s + 2n_s\{(\frac{1}{2n_s})^2 \frac{c_d'}{2}\} = n_s + \frac{c_d'}{4n_s}$$

  which is an EOQ expression in $n_s$[†]. Its minimum over $n_s = 1, 2, 3, \ldots$ is the optimal cost.

[†] 第一项为到达成本，第二项为干线运输成本

- This least cost will always be greater or equal to $\sqrt{c_d'}$ because it is the minimum of $n_s + \frac{c_d'}{4n_s}$ with unrestricted $n_s$ obtained for $n_s^* = (c_d'/4)^{1/2}$.
- Clearly, the underprediction will be most significant when $n_s^*$ is close to an odd multiple of 0.5, or close to zero. We have tested the sensitivity of the EOQ cost expression to errors in the decision variables, which also quantifies this underprediction; as $n_s^*$ increases the underprediction quickly vanishes
- Once $c_d' > 16$ ($n_s^*$ is greater than 2) the difference is below 1%. If $c_d' > 4$ (the value at which $n_s^* = 1$) then the maximum difference stays below 6%. Although for smaller $c_d'$ the difference can grow arbitrarily large as $c_d' \to 0$, that is not the case that is likely to be of interest; the large spacing between terminals recommended by the CA method (much larger than $d_{\max}$) indidates that operating line-haul vehicles is probably an overkill（多此一举）.
- If it were of interest, and a terminal had to be provided, terminal had to be provided, one could force the solution to the CA approach to satisfy the constraint $n_s > 1$. The next section will discuss how more involved constraints can be accommodated within a general CA framework.

- Although exhibiting different errors types, both examples shared a common trait when their errors were largest: the ideal terminal spacing in an interval with constant demand exceeded the length of the interval; i.e., demand varied significantly within the spacing.
- Errors arose because this property violates the stated requirement for the CA approach: $D'(d)$ should vary slowly over distances comparable with $s(d)$. Conversely, the numerical results prove that an error below 1% results if $D(d)$ is piece-wise linear with segments at least three times as long as each $s(d)$.
- Thus, any demand function that can be approximated in this manner should also yield accurate results.

# 目 录 |Outline


1. Other One-Dimensional Location Problems

2. Accuracy of the CA Expression

3. Generalization of the CA Approach

- The CA method can be applied to more complex problems - even problems that defy exact numerical solution. In forthcoming chapters it will be used to locate points in multidimensional (time-space) domains while satisfying decision variable constraints.
- All that is needed is that the input data vary slowly with position, either in one or multiple dimensions, that the total cost can be expressed as a sum of costs over non-overlapping (small) regions of the location domain, and that these component costs (and constraints) depend only on the decisions made in their regions. If this is true, the decomposition principle holds and the CA results approximate the optimal cost accurately.

# Constraints

As a one-dimensional illustration, let us return to the inventory control problem, and let us assume that there is a capacity constraint on shipment size:

$$D(t_i) - D(t_{i-1}) \leqslant v_{\max}$$

This constraint has a local nature because it only involves quantities determined by events close to the time of shipment; i.e., by two neighboring dispatching times and by the amount of consumption between them. For any time $t$ it should be possible to write the constraint approximately as an inequality including only variables and data specific to time $t$.
该约束的物理意义是：两次配送时间的发生的需求量应该小于最大批量

# EOQ formula

Recalling the definition of $H_s(t)$[‡], and using the slow-varying property of $D'(t)$, we can write:

$$D(t_i) - D(t_{i-1}) \approx H_s(t)D'(t) \approx H(t)D'(t)$$

and the constraint can be replaced by the approximation based only on conditions at $t$:

$$H(t)D'(t) \leqslant v_{\max}, \quad \text{or} \quad H(t) \leqslant v_{\max}/D'(t)$$

which must be satisfied for all $t$.
An approximate solution to our problem, thus, is an $H(t)$ that minimizes the total cost subject to this constraint. The optimal $H(t)$ is the least of: (i) $(2c_f/c_iD'(t))^{1/2}$, and (ii) $v_{\max}/D'(t)$. Letting $\Psi\{x\}$ denote the increasing concave function $\{x^{1/2}$ if $x \leqslant 1$; or $[1+x]/2$ if $x > 1\}$, we can express the minimum cost per unit time concisely in terms of the dimensionless quantity, $2c_fD'(t)/(c_iv_{\max}^2)$:

$$c_iv_{\max}\Psi\{2c_fD'(t)/(c_iv_{\max}^2)\}$$

‡ $H_s(t)$ as a step function such that $H_s(t) = t_i - t_{i-1}$ if $t \in I_i$

# Total cost and average cost

Integrated from 0 to $t_{\max}$, this expression approximates the optimal total cost. Note that when the argument of $\Psi$ is less than one, as would happen if $v_{\max}$ is very large, then the expression becomes $[2c_i c_f D'(t)]^{1/2}$. An average cost per item can also be obtained in a similar way; its interpretation as a cost average across items (calculated as if each item was part of a problem with constant conditions, equal to the local conditions for the item) is still valid. In practical cases, a per-item cost estimate can be obtained easily with the following two-step procedure:

- Solve the problem with constant conditions for a representative sample of items and input data,
- Average the solution across all the sampled items to obtain the result.

Note that the cost estimate can be obtained even without defining the decision variables in the first step.

# Practical Considerations

- While for simple problems, such as the one solved above, the solution can be easily automated, more complex situations may benefit from decision support tools with substantial human intervention. The following two-step human/machine procedure is recommended:
  - recognizing that its recommendations may need fine-tuning adjustments, the CA (or other simplified) method is applied to a basic version of the problem without secondary details;
  - trained humans develop implementable solutions that account for the details, perhaps aided by numerical methods that can benefit from the output of the first step.
- In some cases, when time is of the essence humans alone may have to carry out this second step because efficient numerical methods capturing peculiar details may not be readily available, and developing them may be prohibitively time consuming.

- Even without time pressures, if the details are so complex (or so vaguely understood) that they cannot be quantified properly, pursuing automation for the fine-tuning step would seem ill-advised. Fortunately, this is not a serious drawback, significant departures from ideal situations should not increase cost significantly, leaving humans considerable latitude for accommodating details.
- The cost of the two-step procedure (fine-tuned by hand) is compared to the ideal cost without restrictions, and (optionally) to the exact optimal cost obtained with dynamic programming.
- We may find that the fine-tuning step often identifies the exact optimum, and when it does not, the difference between the two-step and the exact optimal costs is measured by a fraction of a percentage point. Furthermore, the two-step and one-step (or ideal) costs are very close; of course, provided that $n_s^*$ is not greater than 50.

# Readings

- Daganzo. Logistics System Analysis. Ch.3. Page 64-73

Any questions?