物流系统分析

Logistics System Analysis

第 10 周　一到多配送问题–异质顾客

One-to-Many Distribution — Strategies for Different Customers

葛乾

西南交通大学　系统所 & 交通工程系

目 录 |Outline

Overview

- Symmetric strategies are extensions from strategies for identical customers
- Asymmetric strategies allow different customer types to be served differently.
- Conditions under which these more complex strategies are likely to be of benefit are also discussed here

目 录 |Outline

Symmetric Strategies

- Let us allow $D_n(t)$ to vary across customers n, and possibly to be non-stationary. With this generalization, even if the demand is stationary, D_n' can vary across n.
- With many customers the individual demand rates should be treated as "details", which we try to avoid. To this end, an expected **demand density rate per unit area** $\lambda(t, \boldsymbol{x})$ is used instead of the specific $D_n(t)$'s.
- $\lambda(t, \boldsymbol{x})$ is assumed to vary slowly with time and location so that the demand in a subregion, $\boldsymbol{P}_p$* of $\boldsymbol{R}$ during a time interval $[t_{m-1}, t_m)$ is:

$$\int_{t_{m-1}}^{t_m} \int_{x \in \boldsymbol{P}_p} \lambda(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x}\boldsymbol{d}t.$$

* $\boldsymbol{P}_p$ is large enough to contain several destinations but of small dimensions relative to $\boldsymbol{R}$

顾客密度与需求不确定性

- Similarly, we define a **customer density**, $\delta(t, \boldsymbol{x})$, which is also allowed to vary with time. Note that we are allowing here for the number and locations of customers to change with time; all we require is that these changes can be approximated with functions $\delta(t, \boldsymbol{x})$ and $\lambda(t, \boldsymbol{x})$, that vary smoothly with t and $\boldsymbol{x}$.
- **Demand uncertainty** is an important phenomenon when **the tours have to be planned before the demand is known at the destinations**. It will be captured by an index of **dispersion function**, as described below.

需求分散度的衡量

- Take a partition $\{\boldsymbol{P}_1, \ldots, \boldsymbol{P}_p, \ldots, \boldsymbol{P}_P\}$ of $\boldsymbol{R}$ and a partition of time into consecutive intervals $\tau_m = [t_{m-1}, t_m)$, and let D_{mp} represent the actual number of items demanded in $\boldsymbol{P}_p$ during τ_m.
- The parameter $\lambda(t, \boldsymbol{x})$ can then be defined as the average demand rate density, so that $\int_{t_{m-1}}^{t_m} \int_{\boldsymbol{x} \in \boldsymbol{P}_p} \lambda(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x} \mathrm{d}t$ now gives the mean of D_{mp}.
- We assume that, for any partition, the variables D_{mp} are independent, and identically distributed. Then their variance can be expressed as:

$$\mathrm{var}\{D_{mp}\} \cong \int_{\tau_m} \int_{\boldsymbol{P}_P} \lambda(t, \boldsymbol{x}) \gamma(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x}\boldsymbol{d}t$$

where $\gamma(t, \boldsymbol{x})$ is an "index of dispersion", with "items" as its physical dimension. $\gamma(t, \boldsymbol{x})$ 是衡量分散度的指标，其物理量纲为数量

需求分散度的衡量 (cont.)

$$\mathrm{var}\{D_{mp}\} \cong \int_{\tau_m} \int_{\boldsymbol{P}_P} \lambda(t, \boldsymbol{x})\gamma(t, \boldsymbol{x})\mathrm{d}\boldsymbol{x}\boldsymbol{d}t$$

- A special case of this model arises if each customer's demand fluctuates independently of other customers, either like a stochastic process with independent increments — such as a compound Poisson process or a Brownian motion process.
- Although in most cases a fixed γ should capture demand fluctuations well, we allow $\gamma(t, \boldsymbol{x})$ to vary slowly with t and $\boldsymbol{x}$. An index equal to zero represents known demand; no uncertainty. This case will be examined next.

物理系统成本构成

Recall the LCF:

$$z = \alpha_0 + \alpha_1 \frac{1}{n_s v} + \alpha_2 \frac{1}{v} + \alpha_3 n_s + \alpha_4 v.$$

where the $\alpha_0, \ldots, \alpha_4$ are the following interpretable cost constants:

$\alpha_0 =$ $(c_s' + c_i r/s + c_i t_s/2)$; handling and fixed pipeline inventory cost per item,

$\alpha_1 =$ $(2rc_d + c_s)$; transportation cost per dispatch,

$\alpha_2 =$ $(c_d k\delta^{-1/2} + c_s)$; transportation cost added by a customer detour,

$\alpha_3 =$ $1/2c_i(k\delta^{-1/2}/s + t_s)$; pipeline inventory cost per item caused by a customer detour and the ensuing stop,

$\alpha_4 =$ c_h/D'; stationary holding cost of holding one item during the time $(1/D')$ between demands.

The constraints are $n_s v \leqslant v_{\max}$ and $n_s \geqslant 1$.

决策变量与目标函数

- For consistency with the literature, we continue to use $H(t, \boldsymbol{x})$ and $A(t, \boldsymbol{x})$ as the decision variables instead of n_s and v. Both formulations are equivalent, since there is a 1:1 correspondence between two sets of variables—the number of stops in a tour is the number of customers in its district, which is given by $n_s \approx \delta(t, \boldsymbol{x})A(t, \boldsymbol{x})$, and the delivery lot size is the consumption during a headway in the area around a customer: $v \approx \lambda(t, \boldsymbol{x})H(t, \boldsymbol{x})/\delta(t, \boldsymbol{x})$.
- Making these substitutions in the LCF and the constraints, and recognizing that $D' = \lambda/\delta$, the cost per item at $(t, \boldsymbol{x})$ can be expressed as:

$$z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$

where α_1, α_2 and α_3 are the constants defined in connection with LCF optimization problem, which now can vary in both time and space; the constraints become: $\lambda AH \leqslant v_{\max}, \delta A \geqslant 1,$ and $\lambda H \leqslant v^{\circ}\delta$.

- The important thing to remember here is not new expression for LCF, but the process followed to derive them and use them. This process is quite general and can be used for problems involving various peculiarities.
- Because it is impossible here to discuss all possible situations, the process is only illustrated with three examples involving stochastic phenomena and requiring some modifications to the equations.
- The first example arises where **items are indivisible and the expected demand per customer per headway is less than one item**;
- The second when **the customer demands are not known until the vehicles make the stop**
- The third when **the vehicles make coordinated adjustments to their routes as demand information becomes known**.

Random Demand: Low Customer Demand

$$\min \quad z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$

$$\text{s.t.}\ \lambda AH \leqslant v_{\max}, \delta A \geqslant 1,\ \text{and}\ \lambda H \leqslant v^{\circ}\delta$$

- It implicitly assume that each customer is visited each time – the number of stops is equal to δA. But if items are indivisible (as opposed to fluids, or very small items) and the demand by individual customers is so low that **some have no demand during a headway, their stops can be skipped**. 低需求率的特点是每次配送不必访问所有顾客。
- For some demand processes, the proportion of stops that can be skipped should decrease with H as $exp^{(-H/H_0)}$, where H_0 is a constant that depends on t and $\boldsymbol{x}$. If the customers in a subregion are alike and their demand is well described by Poisson processes*, then the parameter H_0 is the average time between successive demands at one destination; i.e., $H_0 = D'^{-1} = \delta/\lambda$. For other processes the relationship is similar.

*当需求的变动遵从泊松过程时，需求之间的时间间隔服从指数分布

Random Demand: Low Customer Demand (cont.)

- The effective density of stops is only $\delta[1 - exp^{-H/H_0}]$.
- This expression must be substituted for the parameter δ in the expressions for $(\delta\alpha_2)$ and $(\delta\alpha_3)$ (remember that δ also appears in α_2 and α_3). The optimization and design process can be carried out as described earlier. Although the resulting optimum is slightly more complicated, two extreme cases are quite simple.
- First, if $H \gg H_0$ then the density of stops is δ as before; the solution does not have to be changed. The opposite extreme case with $H \ll H_0$, arising for example if $\delta \to \infty$ but $D' \to 0$, also admits a simple expression for the stop density, even if the demand varies across customers. The expression is $\delta H/H_0 \approx \lambda H$* if items are not demanded in batches; then the number of vehicle stops per tour, $(\lambda H)A$, equals the vehicle load λHA as one might expect.

*每次配送每个顾客被停靠的概率，同样也是每两个配送间隔每个顾客产生的需求量

Random Demand: Uncertain Customer Requests

$$\min \quad z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$
$$\text{s.t. } \lambda AH \leqslant v_{\max}, \delta A \geqslant 1, \text{ and } \lambda H \leqslant v^{\circ}\delta$$

- If α_3 is small* we have seen that the minimum logistics cost will be such that $\lambda AH = v_{\max}$. There is an incentive to dispatch totally full vehicles.
- Let us now see what modifications are needed if the exact demand on a vehicle route is not known accurately when the vehicles are dispatched†.
- The system of interest operates with a **headway** (e.g., daily, weekly, etc.) to be determined, and advertised to customers as a service schedule that is to be met even if the volumes to be carried change with every headway. This scenario can arise for both collection and distribution problems, although for distribution problems of destination-specific items the demand will normally be known.

*It implies that items are cheap

†A case with expensive items is not considered here because if time is of the essence, it is unlikely that one would operate with imperfect information

Random Demand: Uncertain Customer Requests (cont.)

$$\min \quad z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$
$$\text{s.t. } \lambda AH \leqslant v_{\max}, \delta A \geqslant 1, \text{ and } \lambda H \leqslant v^{\circ}\delta$$

- If the size of each delivery v_n is both known and small compared with $v_{\max}$ it should not be difficult to partition the service region into delivery districts of nearly ideal shape with $\sum_n v_n \approx v_{\max}$. Then, the distance formulae hold and the LCF can be used without modification.
- If some delivery lots are comparable to the vehicle's capacity, the routing problem is more difficult because one needs to balance the incentive for filling a vehicle by delivering a right lot size to an out-of-the-way customer with the extra distance that one would have to travel.

Random Demand: Uncertain Customer Requests (cont.)

- In view of the above, our discussion is phrased in terms of collection, although hypothetical distribution problems with uncertain demand would be mathematically analogous.
- For collection problems some of the vehicles may be filled before completing their routes, which would cause some of the demands to go unfulfilled.

Random Demand: Uncertain Customer Requests (cont.)

- The overflow customers (still needing visits) could be covered in the same headway by collection vehicles with unused cargo space or, failing that, by vehicles dispatched from the depot.
- Clearly, if some vehicles can be rerouted before returning to the depot, some distance can be saved. Dynamic routing introduces modeling complexities that will be discussed later. For now we assume that **all the overflow customers are visited by a separate set of secondary vehicle routes based at the depot and planned with full information**.
- This information is gathered by the original (primary) vehicles, which are assumed to visit all the customers. Because items are “cheap”, secondary vehicles should also travel full.

Random Demand: Uncertain Customer Requests (cont.)

- The decision variables are A and H, as before, but now the capacity constraint must be replaced by an **overflow cost** which depends on A and H. A new trade-off becomes clear.
- If the average demand for a tour satisfies $\lambda AH \ll v_{\max}$, then the overflow cost will be negligible, but most primary vehicles will travel nearly empty.
- On the other hand, if $\lambda AH \approx v_{\max}$, a larger number of customers will overflow on average—the actual number will depend on the variability of demand as captured by its index of dispersion, γ.

Random Demand: Uncertain Customer Requests (cont.)

$$\min \quad z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$
$$\text{s.t. } \lambda AH \leqslant v_{\max}, \delta A \geqslant 1, \text{ and } \lambda H \leqslant v^\circ \delta$$

- Instead of a total cost per item, we work with a **cost per unit time and per unit area**. For given A and H, the transportation cost per unit time and unit area for primary tours is approximately independent of the overflow; it is well approximated by the product of the constant factor λ, and the first two terms of z:

$$\frac{\alpha_1}{AH} + \frac{\delta\alpha_2}{H}$$

  Strictly speaking, this expression is an upper bound because it **ignores the local delivery distance** that it is saved by the stops that are skipped.

Random Demand: Uncertain Customer Requests (cont.)

- Note that, especially when the fraction of tours overflowing is small, the **overflow customers will tend to be geographically distributed in widely spaced clusters of customers corresponding to overflowing tours**. Because the overflow transportation cost formulas with clustered destinations are more complicated, two simple bounds will be used instead to approximate the secondary distance traveled*. 超量顾客的分布更稀疏
- It should be intuitive without a formal derivation that **smearing the clusters uniformly over R** increases the distance traveled, while **collapsing them into a single point** decreases it. Upper and lower bounds for secondary distance are derived below, imagining that clusters are either spread or fused in this manner.

*Blumenfeld and Beckmann, 1984, have developed formulas for VRP's with clustered demand points

Random Demand: Uncertain Customer Requests (cont.)

辅助运输发生的次数

- An expression for f_0, the fraction of items that must be delivered or collected as overflow, will be derived shortly.
- Assume for now that f_0 is given. Then the number of secondary (overflow) tours per unit area is $\lambda H f_0 / v_{max}$*, and the number of stops is close to $f_0 \delta$.
- This expression implies that **the fraction of items overflowing is the same as the fraction of customers**; the expression is exact if primary vehicles don't deliver (or collect) partial lots, and is also a good approximation in other cases.

*λH 是区域内的需求量，除以 v_{max} 表示主要配送所服务的顾客数，再乘以 f_0 表示辅助配送所服务的顾客数

Random Demand: Uncertain Customer Requests (cont.)

超量顾客分散/集中分布时，辅助运输的距离

- With **de-clustered overflowing customers**, the upper bound to the secondary distance per unit area is thus:

$$\frac{2r\lambda H f_0}{v_{\max}} + k(f_0\delta)^{1/2}.$$

  We are assuming here that the total number of customers is greater than the squared number of stops per vehicle: $Nf_0 \gg (v_{\max}\delta/\lambda H)^2$

- With **perfectly clustered groups** the density of stops equals the density of incomplete primary **tours**. If we let g_0 denote the probability that a tour overflows, then this density is g_0/A; thus a lower bound for the distance per unit area is:

$$\frac{2r\lambda H f_0}{v_{\max}} + k(g_0/A)^{1/2}.$$

- The secondary transportation cost per unit area and unit time is obtained by multiplying either distance bound by c_d/H, and adding to the result the cost of stopping. For the upper bound we have:

$$\begin{aligned}\text{overflow transport cost} \approx & c_d\left[\frac{2r\lambda f_0}{v_{\max}} + \frac{k(f_0\delta)^{1/2}}{H}\right] + c_s\left(\frac{\lambda f_0}{v_{\max}} + \frac{f_0\delta}{H}\right) \\ = & \alpha_1(\frac{\lambda f_0}{v_{\max}}) + \frac{k(f_0\delta)^{1/2}c_d}{H} + \frac{f_0\delta c_s}{H}\end{aligned}$$

- For the lower bound, the factor $(f_0\delta)^{1/2}$ of the second term should be replaced by $(g_0/A)^{1/2}$. If the overflow is so small that only a few secondary tours are used, $Nf_0 < [v_{\max}\delta/\lambda H]^2$, then k should be replaced by k' and r should be set to 0, regardless of position.

- Either on primary or secondary tours, items reach the destination at regular intervals, as required, approximately H time units apart. Thus, the stationary holding cost per unit time and unit area is:

$$\text{holding cost} \approx c_h(\lambda H)$$

- We are now ready to write the logistic cost function for our problem. In practical situations one would expect the difference between the upper and lower bound to be small. Therefore, we will use one of these bounds (the upper bound) below.

Random Demand: Uncertain Customer Requests (cont.)

- In terms of total cost per unit time and unit area (the sum of primary and secondary transportation costs, plus the holding cost), the upper bound is

$$\lambda z \cong \frac{(\alpha_1)}{AH} + \frac{(\delta\alpha_2)}{H} + (\alpha_1 \frac{\lambda}{v_{\max}}) f_0 + (k\delta^{1/2} c_d) \frac{f_0^{1/2}}{H} + (\delta c_s) \frac{f_0}{H} + (\lambda c_h) H,$$

  where the parenthetical items are constants and the rest $(A, H,$ and $f_0)$ are decision variables. Note that the constant handling cost, α_0, has been omitted from the LCF.

Random Demand: Uncertain Customer Requests (cont.)

- The fraction of items that overflow is related to A and H.
- Recall the following equations

$$\text{mean}\{D_{mp}\} \cong \int_{t=t_{m-1}}^{t_m} \int_{x \in \boldsymbol{P}_p} \lambda(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x} \mathrm{d}t.$$

$$\text{var}\{D_{mp}\} \cong \int_{\tau_m} \int_{\boldsymbol{P}_P} \lambda(t, \boldsymbol{x}) \gamma(t, \boldsymbol{x}) \mathrm{d}\boldsymbol{x} \mathrm{d}t$$

The mean and variance of the number of items to be carried by a primary vehicle are λAH and $\lambda A \gamma H$. The expectation of the excess of this random variable over $v_{\max}$ is the average overflow for the vehicle.

Assuming that the demand is approximately normally distributed, and letting Φ denote the standard normal CDF (and Φ' its derivative —the PDF), we can therefore write:

$$\begin{aligned} f_0 \cong& \frac{1}{\lambda AH} \int_{v_{\max}}^{\infty} (x - v_{\max}) \mathrm{d}\Phi \left(\frac{x - \lambda AH}{(\lambda A \gamma H)^{1/2}} \right) \\ =& (\lambda AH/\gamma)^{-1/2} \Psi \left[\frac{(\lambda AH - v_{\max})}{(\lambda A \gamma H)^{1/2}} \right] \end{aligned}$$

where

$$\Psi(z) = \int_{-\infty}^{z} \Phi(w) \mathrm{d}w = \Phi'(z) + z\Phi(z)$$

which is a convex function increasing from zero (when $z \to -\infty$) to ∞ (when $z \to \infty$). Note that f_0 may depend on position and time.
本式计算的实际上是：假设配送量服从正态分布时，其取值大于等于于 $v_{\max}$ 的概率。该分布的参数为 λ, A, H, γ。

$$\lambda z \cong \frac{(\alpha_1)}{AH} + \frac{(\delta\alpha_2)}{H} + (\alpha_1 \frac{\lambda}{v_{\max}}) f_0 + (k\delta^{1/2} c_d) \frac{f_0^{1/2}}{H} + (\delta c_s) \frac{f_0}{H} + (\lambda c_h) H,$$

- Thus, λz should be minimized, subject to the expression of f_0.
- The procedure is simple. Conditional on AH, i.e. on the average vehicle load per district, f_0 is fixed and λz only depends on H; the optimal headway can be obtained in closed form from the expression of λz as an EOQ trade-off involving the 2nd, 4th, 5th and 6th terms of that expression. The resulting cost is only a function of AH, which can be minimized numerically.
- The procedure also works for the lower bound, and when the number of secondary tours is low. For the lower bound one should replace the fourth term of λz by $kc_d(g_0/A)^{1/2}/H$, where $g_0 = \Phi(z)$. Note that g_0 is fixed if AH is fixed, like f_0.

- Cost estimates and guidelines for the construction of a detailed strategy can be obtained as usual, by repeating the minimization for a few combinations of $(t, \boldsymbol{x})$.
- We could also verify that the final strategy and the resulting cost do not change much if the overflow local distance term is replaced by the lower bound.

Dynamic Response to Uncertainty

很多情况下，车辆的路径可以动态调整

- In many applications, vehicle routes can be adjusted dynamically during the course of operation. For example if a collection truck of an express package carrier falls behind schedule, central dispatch can reassign some of its remaining customers to currently underutilized trucks. 当某辆用于收集快件的卡车滞后于时刻表时，配送部门可将其剩余顾客分配给目前未被占用的卡车
- If a firm can do this systematically with an efficient control strategy, it should be able to operate with fewer vehicles.

决策尺度

- To design such a system we must make a single set of **planning*** **decisions** at the beginning of the planning period, e.g., choosing **# trucks**; and then a stream of **control decisions** that change dynamically as information is revealed over time.

*or configuration

规划阶段的策略要求

To minimize the combination of fixed and operating costs, configuration decisions must anticipate and accommodate the long-run needs of the control strategy; that is, the system should be planned for control. This is difficult to do exactly but can be achieved approximately if we can find a family of control strategies that is:

1. parametrizable (describable in terms of **just a few parameters**);
2. appealing (containing a **near-optimal strategy** for the configuration of **every reasonable system**);
3. simple (with a **predictable expected cost**).

Properties (1) and (3) guarantee we can write an LCF that captures approximately all fixed and recurring costs in terms of the configuration variables and control parameters. Property (2) guarantees that good control parameters exist for every reasonable configuration.

规划阶段的策略要求 (cont.)

- Hence, the minimum of the LCF is an “appealing” plan. Since an analytic expression exists the minimum can be searched effectively with conventional optimization methods, even if the number of variables and parameters is considerable.
- The selection of a proper family is more an art than a science. The temptation is always to look for the most efficient control strategies, excelling at (2), even if they fail the simplicity test (3). The problem with this approach is that a search for the optimum configuration cannot then easily incorporate the effects of control. The result can be gross sub-optimization.
- For planning purposes we prefer to look for **idealized*** **control strategies** that can be systematically analyzed. This allows us to explore a much larger solution space when configuring the system. The idealized strategies play the role of approximations to the more refined strategies during the optimization process, but the refined strategies can still be used when the system is operated.

*less efficient

配送区域的划分与决策变量

- Let us consider again the load-constrained system, but assume now that $H = 1$ day as in package collection systems. We want to configure a system where vehicles that are partially filled at the end of their runs can cover the overflow customers of other vehicles. Although very complex dynamic routing strategies can be designed to achieve this goal, we shall be satisfied with a simple one that is obviously sub-optimal but improves significantly on the static approach
- We partition the service region into an **inner region close to the depot** (region 2) and an **outer fringe** (region 1). Only customers in region 1 are allocated to primary tours. We use only one planning variable: **# of primary service zones in region 1**, which equals the number of vehicles m. The **radius of the inner region**, r_T, is our **control parameter**.

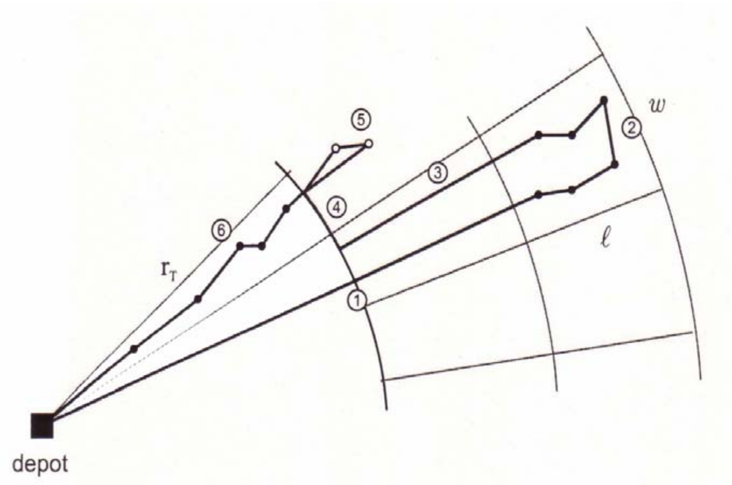

Figure: The idealized control strategy has two phases with several steps

规划阶段的服务过程

- In phase one vehicles travel to their service zones (step 1), serve their customers (step 2), and either return to the depot, if filled, or else stop at the boundary between regions 1 and 2 (step 3). Unfilled vehicles wait there for the start of the second phase, until all vehicles are done.
- Then, they are repositioned along the boundary in anticipation of serving carefully designed groups of remaining customers (step 4). The size of these groups is chosen to be consistent with each vehicle's available capacity. Vehicles first serve the part of their group in region 1(step 5), then the part in region 2 (step 6). Region 2 customers are arranged in wedges that can be served efficiently as vehicles return to the depot. Finally, if any customers remain unserved, they are served with a set of secondary tours (step 7). Note that virtually no customers require such secondary tours when systems are configured optimally.

规划阶段的服务过程 (cont.)

- This strategy generalizes the static procedure, since the effects of the latter can be essentially achieved by setting $r_T = 0$. Although the new strategy is sub-optimal, it has clear efficiencies over the static procedure; thus, it is "appealing" in the sense of (ii). The strategy also has properties (i) and (iii), since it* is parameterized by the inner radius r_T and is simple.
- An analytic approximation for the LCF is given in Erera (2000). The approximations in this reference were designed to be most accurate for intermediate values of r_T, where the optimum was expected to be. The formulae are not given here because they would take too long to explain, but the qualitative results are interesting.

*注意课本 142 页最后一段倒数第五行 t 应为 it

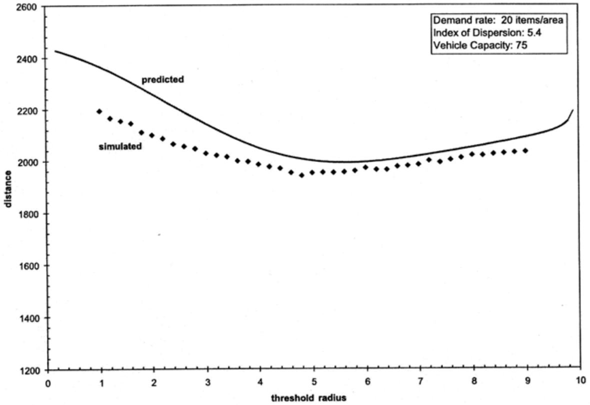

Figure: This figure shows how the approximate total distance per day varies as a function of r_T for a test problem, after the number of vehicles m was optimized. The figure also includes a dotted line from a simulation that used the recommended values of m and r_T, and a more sophisticated control algorithm. This curve gives the actual distance that could be expected in an implementation.

- Reassuringly, the value of r_T recommended by the optimization (the minimum of the solid line) yields a near-minimum actual distance. Note from the figure that this distance is considerably smaller than that achieved with the static strategy ($r_T = 0$).
- Erera (2000) shows with a battery of 20 problems that the reduction in the required number of vehicles is even greater.
- The portion of the vehicle fleet required by uncertainty (the "fleet penalty" in Erera's lingo) was reduced by 50% or more in 19 out of 20 cases and by more than 70% in half of the cases. The median reduction in the "distance penalty" due to uncertainty, on the other hand was only about 30%.

目录 |Outline

The scenario

- We now explore the advantages of offering different service levels to customers with **different consumption rates** and/or **different holding costs**.
- Because these differences are likely to be most notable for collection problems, our discussion will be phrased in these terms — **factories and manufacturing plants typically consume a wide selection of parts and raw materials even if their product line is homogeneous**.
- Before explaining how asymmetric collection strategies can be designed, we introduce why they are desirable with a very simple example with two customer types.

The LCF

$$\min \quad z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + \delta\alpha_3 A + c_h H + \alpha_0$$
$$\text{s.t. } \lambda AH \leqslant v_{\max}, \delta A \geqslant 1, \text{ and } \lambda H \leqslant v^\circ \delta$$

- Consider a problem with stationary conditions (i.e. λ and δ independent of time) obeying the LCF for which it is desirable to fill the vehicles. More specifically, we assume that: (i) the third (**pipeline inventory**) term of LCF **can be neglected** because items are "cheap", and (ii) that only constraint $\lambda AH \leqslant v_{\max}$ plays a role because **storage room at the origins is plentiful** and the customer density is so large that **the ideal # of vehicle stops is sure to exceed 1**. We also assume that the stop cost c_s can be neglected.

The LCF (cont.)

$$\min z = \frac{\alpha_1}{A\lambda H} + \frac{\delta\alpha_2}{\lambda H} + c_h H + \alpha_0 \text{ s.t. } \lambda A H \leqslant v_{\max}$$

- Let us now examine how the optimal system cost depends on λ and δ. Because z decreases with A for any H, its minimum is reached for as large a district area A as possible.
- Therefore, as expected, the vehicle capacity constraint must hold strictly: $A = v_{\max}/(\lambda H)$. On making this substitution and minimizing the resulting EOQ expression with respect to H, a simple formula for the cost per item z^*, is obtained.

- If α_4 $(c_h H = \alpha_4 v)$* is replaced by its expression in terms of δ and λ (i.e., $\alpha_4 = c_h \delta/\lambda$), and the result is expressed in cost units per unit time and unit area, the formula becomes:

$$\lambda z^* = \beta_1 \lambda + (\beta_2 \lambda)^{1/2} \delta^{1/4}.$$

where $\beta_1 = \alpha_0 + \alpha_1 / v_{\max}$ and $\beta_2 = 4 c_h c_d k$. Notice that λz^* increases at a decreasing rate with λ, δ and β_2; this concavity encourages discrimination

*Recall $\alpha_4 = c_h / D' = c_h H / v$.
*注意课本 144 页公式 4.27a 笔误

Two different types of customers

- Suppose that there are two customer types, $n = 1, 2$, with demand characteristics (λ_n, δ_n) and with different c_h, so that β_2 is different for the two customer types: $\beta_2^{(1)}$ and $\beta_2^{(2)}$. (We use n to index customer classes, instead of customers.)
- Note then that $\lambda = \lambda_1 + \lambda_2$ and $\delta = \delta_1 + \delta_2$.

Separate delivery

If the two customer classes are treated completely separately, as if the other did not exist, the combined cost per unit time and unit area, instead of being given by $\lambda z^* = \beta_1 \lambda + (\beta_2 \lambda^{1/2})\delta^{1/4}$, would be:

$$\lambda z^* = \sum_{n=1}^{2} \left[\lambda_n \beta_1 + \left(\beta_2^{(n)} \lambda_n \right)^{1/2} \delta_n^{1/4} \right]$$

$$= \lambda \beta_1 + \sum_{n=1}^{2} \left(\beta_2^{(n)} \lambda_n \right)^{1/2} \delta_n^{1/4}.$$

When this strategy is best?

It is best if:

$$\sum_{n=1}^{2}(\beta_2^{(n)}\lambda_n)^{1/2}\delta_n^{1/4} < (\sum_{n=1}^{2}\beta_2^{(n)}\lambda_n)^{1/2}(\sum_{n=1}^{2}\delta_n)^{1/4} \quad \text{Cauchy-Schwarz Inequality}$$

If the two customer types are similar, this inequality does not hold. Therefore, a symmetric strategy is best: items should be shipped together because with the higher demand density resulting from amalgamation vehicle tours can cover smaller zones and save operating costs. This is not always the case, however.

*The inequality holds when $\beta_2^{(1)}\lambda_1/\delta_1^{1/2} \neq \beta_2^{(2)}\lambda_2/\delta_2^{1/2}$

When this strategy is best?

$$\sum_{n=1}^{2}(\beta_2^{(n)}\lambda_n)^{1/2}\delta_n^{1/4} < (\sum_{n=1}^{2}\beta_2^{(n)}\lambda_n)^{1/2}(\sum_{n=1}^{2}\delta_n)^{1/4}$$

- This will hold if one set of suppliers is highly concentrated $\delta_1 \approx 0$ while producing many items that are expensive to store ($\lambda_1\beta_2^{(1)}$ large), and the other set has opposite characteristics (δ_2 is large but $\beta_2^{(2)} \approx 0$). 第一类供应商集中且存储费用高；第二类供应商更分散且存储费用更低
- Separate service for the two sets is then reasonable because the distribution strategies for both sets should be different. For the second set one would like to save operating costs at the expense of holding cost (one would use a large H in order to reduce the area served by each vehicle) and for the first set one would do the opposite. 第二类供应商应该用更低频率配送，以通过适量提高存储费用节省运输费用，另一类反之。

Combined delivery v.s. separate delivery

- In both cases the local operating costs plus the holding cost $((\beta_2^{(n)}\lambda_n)^{1/2}\delta_n^{1/4})$ would be close to zero. However, if both items types are combined together, neither of the factors on the right side $((\sum_{n=1}^{2}\beta_2^{(n)}\lambda_n)^{1/2}(\sum_{n=1}^{2}\delta_n)^{1/4})$ of is close to zero — service has to be moderately frequent because some of the items are expensive to store, and tours must cover moderate size areas because all destinations have to be visited.
- Clearly, the requirements of the two sets of customers interfere with each other, increasing cost dramatically.

现实世界的情况

- This phenomenon explains why **separate logistic systems are used to carry widely different items in real life**, even if from a transportation standpoint alone it would seem wise to combine them.
- It should not be surprising to find several transportation modes (taxis, limousines, buses, etc.) at the disposal of passengers exiting an airport. For freight transportation, the differences in the requirements of various customers are less likely to merit discriminating service; but the possibility should be considered.

差异化策略

- For general problems, the example just described suggests that cost may be reduced if the set of all customers is divided into classes with different characteristics, served with separate collection systems .
- For a given set of classes, total cost can be easily estimated — the cost and structure of near-optimal symmetric strategies would be used within each of our subsystems.
- The tricky part is **defining the customer subsets that will minimize total cost**. Daganzo (1985) presents a simple dynamic programming procedure to achieve this goal without detailed customer information — the method only uses the frequency (probability) distribution of customer characteristics — and shows in the process that the optimal solution would rarely exhibit more than 2 or 3 classes. When it is found that cost is minimized with only one class, discriminatory service is not cost-effective.

何时有效?

- Although we have ignored the pipeline inventory cost in this lecture, and have also assumed that the same transportation mode is used for all the subsystems, this is not a prerequisite for discriminatory service to be attractive.
- It is impossible to discuss here all the possible cases that can arise in detail, but a general statement can be made: if customers are very different, then we should check **whether dividing them into a few classes with (highly) different characteristics** — and serving them separately — can reduce cost; this is unlikely to result in much gain when customers are not very different, though.

不同配送频率

- With the approach just described, each customer class n is designed separately and is characterized by design parameters A_n and H_n.
- By restricting these design parameters somewhat, Hall (1985) has developed a strategy that **allows customers from all classes to share the transportation fleet** while **being visited at different frequencies**. He requires A to be the same for all customers and each H_n to be an integer multiple of the time between dispatches H; that is, $H_n = m_n H$, for an integer m_n. He assumes that vehicles are dispatched at times $t = 0, H, 2H$, etc., visiting each time $(1/m_n)$th of the customers in every class n. This allows the effective stop density, $\sum_n \{\delta_n / m_n\}$, to be greater than for any class alone while ensuring that individual customers are only visited every m_n dispatches; it decreases the local transportation cost.

单个顾客的不同策略

- With the help of f_0, a variable denoting the fraction of customers served in each period, Hall's strategy can be defined without resorting to classes.
- Accordingly, the symbol "n" now reverts to its original meaning, indexing individual customers. We seek the optimal m_n for individual customers, as well as the optimal H and f_0. As done at the outset, let us assume that the conditions are such that vehicles will be dispatched full.

- Then, the line-haul motion cost per item is $\alpha_1/v_{\max}$, and does not depend on the allocation scheme for customers. The local motion cost per unit time and unit area is:

$$c_d k \frac{(f_0\delta)^{1/2}}{H} + c_s f^0 \frac{\delta}{H}.$$

- This somewhat conservative estimate assumes that stops are randomly and uniformly distributed within subregions of $\boldsymbol{R}$ larger than a collection district; it may be on the high side if customers of a similar kind cluster together.
- The holding cost per unit time in a subregion of unit are $\boldsymbol{P}$ is:

$$\sum_{n\in\boldsymbol{P}} c_h^{(n)}(m_n H)D'_n.$$

分解方法

- The system can be designed with a simple decomposition method. Conditional on f^0 and H, the local motion cost is fixed; thus, cost is minimized by the m_n's that minimize the holding cost. These m_n's, to be consistent with f^0, must satisfy:
$$\sum_{n \in \boldsymbol{P}} 1/m_n = f^0 \delta.$$
Once the m_n have been found, the conditional total cost is obtained. Testing various values of f^0 and H, we can identify a near-optimal solution.
- Alternatively, if one replaces the constraint $[m_n = 1, 2, 3, \ldots]$ by $[m_n > 1]$, a simple approximation for the minimal holding cost for a given f^0 and H can be obtained. The optimal strategy is then defined by the minimum over f^0 and H of the sum of this approximation and the local motion cost expression.

目录 |Outline

One of the reasons for the very extensive literature on algorithms to vehicle routing problems is that in actual applications almost every problem has some peculiarity that renders it unique. We have already seen that there can be a variety of cases depending on:

1. the relative size of the number of tours and the maximum number of stops per tour.
2. the relative cost of rent, inventory, and operating costs.
3. limitations to route length and storage space
4. dissimilarity in the values of items and the demand rates at different destinations
5. amount of uncertainty as to the customer lot sizes.

In addition (and this is not an exhaustive list) one might find situations in which time enters the problem because customers request service during certain “time windows”, or there is a limit to the amount of time an item can spend in transit (perishable items). There also are situations where vehicles do both distribution and collection (routing with backhauls), and situations where vehicle loading considerations make it advantageous to visit customers in an order which does not minimize the total distance traveled.

- At the core of our proposed two-step method for solving general distribution problems there should be a simple and efficient routing algorithm, whose performance can be quantified by means of simple formulas using average density as an input, instead of detailed customer locations. It is then a simple matter to add holding and pipeline inventory costs to the motion cost to define a logistic cost function. If routing/scheduling strategies can be defined in terms of a few decisions variables that are constrained only locally in the time-space domain, then the minimum of the (constrained) logistic cost function will approximate the cost generated by items in different portions of the time-space domain. The CA approach can be used.

- Some routing cost models that allow this to be accomplished already exist. They are now briefly reviewed. Simple transportation cost formulas have been proposed for time-window problems (Daganzo, 1987a,b). The results show how cost increases with the narrowness of the windows, and with the proportion of customers with tight requirements. The proposed routing strategy uses a different set of delivery districts for the customers in each time window, and staggers the zones in such a way so as to leave most vehicles in favorable locations at the beginning of each new window period.

- Perishable items such as newspapers (Han, 1984, and Han and Daganzo, 1986), lead to VRP structures which are similar to those arising from the vehicle route length limitations discussed in Sec. 4.4.1. The main difference is that service districts that are far away form the depot should be (i) more elongated than usual and (ii) covered in a one-way pass that begins at the end of the district that is close to the depot and terminates at the far end. Although this modification increases the line-haul distance traveled, it also allows distribution to begin sooner and the districts to include more stops.

- Models with both pick-ups and deliveries have been constructed for public transportation systems (Daganzo, Hendrickson and Wilson, 1977, Hendrickson, 1978) serving one focal point and a surrounding area. The strategies examined in these early works, however, are not as general as possible; they only consider two extreme cases for a partition of the surrounding area into service zones. More recently, Daganzo and Hall (1990) present an improved cost model for routing with backhauls, emphasizing cases where the total flow in one direction (e.g. outbound from the depot) is a few times larger than in the other direction.

- The basic idea is briefly summarized below for the case where the dominant flow is outbound; the reverse situation is similar. One simply constructs distribution tours as if there were no pickups, allocates each pickup to the nearest return leg of a distribution trip (or "spoke"), and finally modifies the vehicle tours in recognition of the newly assigned stops. Because the density of spokes increases rapidly toward the depot, significant vehicle deviations are only required for pickups near the outer fringe of the region. Pickup miles on the fringe can be reduced by ending the outermost delivery tours at the far end of their districts and by other modifications that are geared to optimize the spatial distribution of spokes. In fact, it is shown in Daganzo and Hall (1990) that under some conditions it is almost as if the secondary stops added only a stop cost and no distance cost. Hall (1993) has applied the concept of spokes to the VRP problem for deliveries only, in which customers demand large and small items.

- Another complication that deserves attention involves the interaction of vehicle loading and routing. When items have awkward shapes and are large, so that only a few fit in a vehicle, vmax may not be fixed; it may depend on the specific customers that are visited or even the order in which they are visited. The latter phenomenon may arise if weight distribution restrictions, for example, dictate that some items (and thus some stops) must be handled before others. This topic is very complex and hard to handle generally; see Hall (1989) and Ball et al. (1995a) for example.

Another area where further results may be desirable involves the interaction of physical distribution with production schedules. This interaction sometimes offers an opportunity for further cost reductions.

- This subject was broached in Sec. 4.3.3 (Inventory at the origin), where it was suggested that production of (destination-specific) items should be rotated among geographical customer regions every headway H. Dispatching the vehicles to a region immediately after its production run was completed greatly reduced the holding costs at the origin. It was assumed that production would be coordinated with transportation in this manner without much of a penalty.
- More likely, though, there may be a set-up cost associated with each switch in production item types. In this case production costs may be reduced by switching less frequently and holding higher inventories at the origin. An integrated solution can then be obtained by including in the logistic cost function the production set-up costs, e.g., as explained below.

- If no attempt is made to coordinate the production schedule with the physical distribution schedule, then the inventory at the origin of items of a certain type can be decomposed as shown in Figure into a (shaded) component which depends on the time between setups for that item type, H_s , and a (dotted) component which depends on the transportation headway, H:

$$\text{average inventory cost per item at origin} \approx \frac{c_i}{2} + \frac{c_i}{2} H.$$

We are assuming that the number of item types is large and, therefore, the steps of the production curve are nearly vertical. Similar conclusions can be reached for few item types.

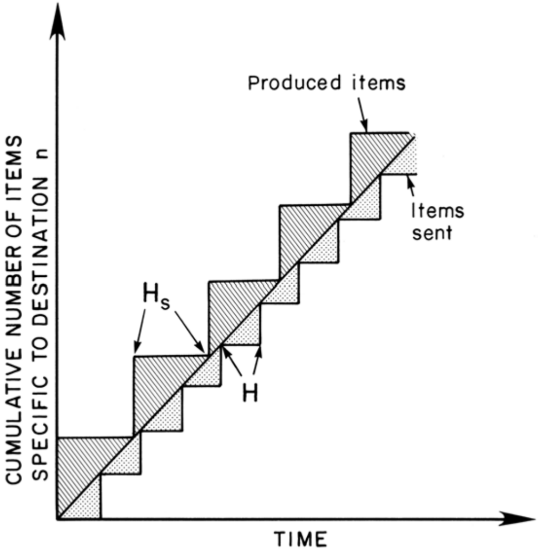

Figure: Inventory accumulation when no attempt is made to coordinate production and distribution

- The maximum accumulation also decomposes in a similar manner:

$$\text{maximum accumulation} \approx H_s D' + HD$$

- Because production costs depend on H_s and not on H, the sum of the production and logistics costs is made up of two components: (i) a production component with only production decision variables (including H_s), and (ii) a logistic component with only logistics variables (including A and H). Logistics and production decisions, thus, can be made independently of each other.

By selecting H to be an integer submultiple of H_s, or vice versa, it is possible to reduce the inventory time at the origin by an amount equal to the smallest of H and H_s, and the maximum accumulation becomes the difference between the maximum and the minimum of $H_s D'$ and HD'.

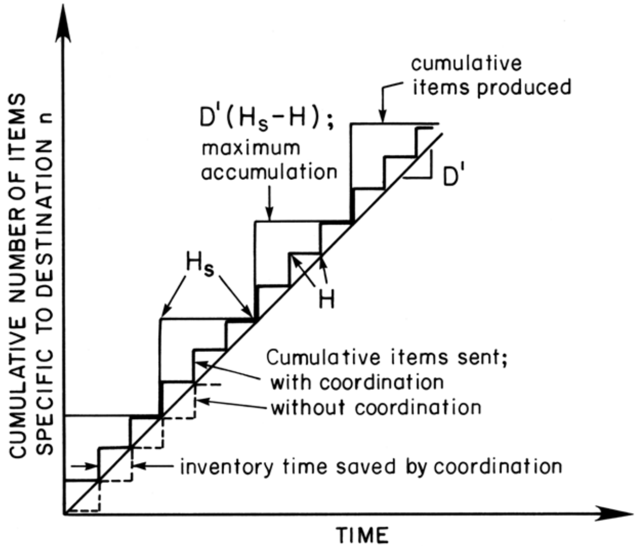

Figure: Inventory accumulation with coordinated schedules with $H_s = 3H$

- If this kind of coordination is feasible, the sum of the production and logistics costs no longer decomposes, and a coordinated production and distribution scheme should be considered.
- Blumenfeld et. al. (1985a) and (1986) have examined the case where each district is constrained to contain only one destination and all shipments are direct ($n_s = 1$). They illustrated situations where coordination of production and distribution is most conducive to cost savings, and provided a bound on the maximum possible benefit.
- Further research may be worthwhile to relax the $n_s = 1$ assumption and to allow more destinations than item types.

- Throughout this talk it was assumed that the total production rate* could be adapted to the changing demand without penalty. In practice, though, this is rarely so, even if the items produced are generic. (It is more costly to change the quantity of items produced than the kind of items produced because to adjust the production rate one needs to hire extra labor, pay overtime or fire labor as needed — and the penalty for these actions is large; Newell, 1990, has examined the production rate adjustment process.) To conclude this lecture, we show that this seemingly strong assumption can often be relaxed.

*not just the schedule by item type

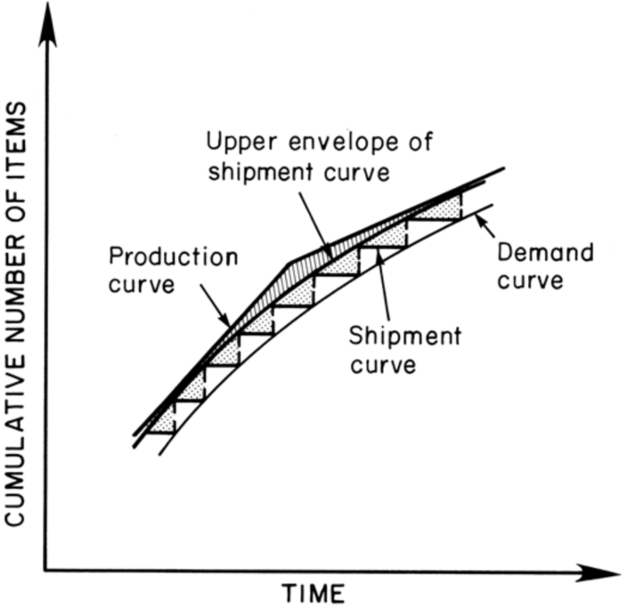

Figure: Production for a gradually decreasing demand

- The figure shows how a production curve may be adapted to a gradually decreasing demand; the objective is tracking the smooth envelope to the crests of the shipment curve (which varies like the demand curve) as closely as possible, without many production rate changes. We had already known that for a similar model described previously, lot size decisions were independent of production decisions; fortunately, this is also true now.
- In this figure, the inventory at the origin decomposes in two components: (i) a (shaded) component, which is due to the discreteness in the production rate changes and is independent of the shipping schedule, and (ii) a dotted component which is the same as if the production schedule was adjusted continuously as assumed in this chapter. Thus, costs can be divided into two components affected respectively only by production, or only by logistics decision variables.

Any questions?

Readings

- Daganzo. Logistics System Analysis. Ch.4. Page 133-153.