Logistics and Supply Chain Review

Tu Ni

AY2019/2020 Semester 1

1 Introduction

1.1 What's supply chain

A supply chain consists of the activities and infrastructure whose purpose is to move products from where they are produced to where they are consumed. Further, supply chain management is the set of practices required to perform the functions of a supply chain and to make them more efficient, less costly, and more profitable. Normally, we can represent a supply chain as a schematic network:
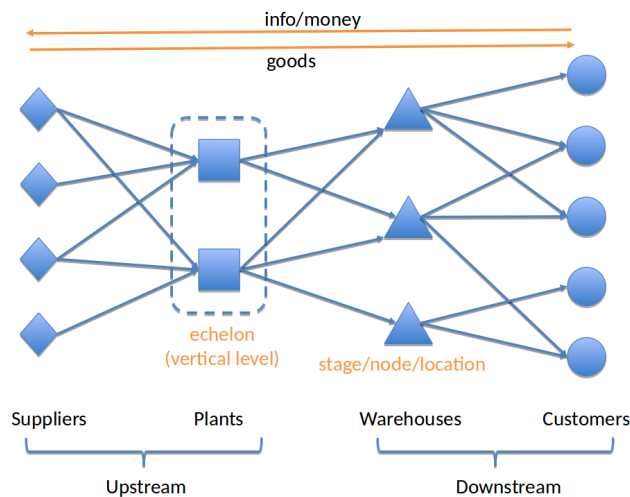

Figure 1: Supply Chain

The structure of supply chain network can vary from serial, converging, diverging, to general tree structure, in which intermediate stages are playing importance roles. The benefits we can get from this kind of structure contain three main aspects:

- Economies of Scale: it's cost-efficient for production or transportation by aggregating demand.

- Risk Pooling: the hyper-connected structure enhances the robustness when facing variations/uncertainty of supply or demand.

- Cross-docking: goods can be sorted or re-organized at intermediate nodes to reduce cost.

Based on the chain/network structure, we can observe physical flow from upstream to downstream and information/money flow in the other way around. For the physical flow, we need to design the way of delivering goods and optimize over it, wherein VRP is an essential model; For the information flow, we need to understand the difference between centralized system and decentralized one.

1.2 What to optimize in supply chain

Since supply chain network is quite complex, we are not able to(and no need to) globally optimize everything together for lower cost, higher profit and better customer value. Instead, we usually focus on one or two key elements of supply chain with single objective.

One perspective is customer value, which includes four major topics:

- Quality: a good supply chain should finally provide high-quality product or service for customers.

- Availability: make-to-stock starts from supply(inventory) while make-to-order deals with demand.

- Assortment: different kinds of product combinations to attract customers.

- Flexibility: whether customer demand can be fulfilled under uncertainty.

The other perspective is about the trade-off between cost and benefit. Mathematically, we can classify cost structure as linear or non-linear cost

$$f(x) = cx; \quad f(x) = \sqrt{x}$$
$$f(x) = \sum_{i=1}^{N} c_i x_i; \quad f(x_1, x_2) = x_1 x_2$$

fixed or variable cost

$$c(x) = \begin{cases} f + cx, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

and even step function. Categorically, we have three types of cost:

- Facility-related Cost: capacity, location...

- Transportation Cost: routing...

- Inventory Cost: ordering, holding, shortage...

1.3 What is difficult in supply chain

The first difficulty comes from uncertainty, which includes supply, demand, price, exchange rate, production yield, competition and future information. Basically we want to study how supply chain should be designed and operated for different kinds of uncertainty.

The second difficulty comes from modeling. The practical situation in supply chain is hard to characterized exactly, but we care about those models with sufficient characteristics that can result in useful insights. Mathematically, if optimal solution can not be found, we may also get good feasible solution with rigorous bounds or performance guarantee to deliver key message.

1.4 How to optimize in supply chain

We can optimize three levels of decisions in supply chain

- Strategic: takes effect over a long time horizon such as location and capacity of warehouses/facilities.

- Tactical: take effect over a moderate time such as assignment and resource allocation.

- Operational: take effect over a short time such as real-time matching and routing.

1.5 Where to apply techniques in supply chain

Beyond traditional supply chain problems, many application topics are also developed for extension:

- Sustainable OM/SC: environmental impacts, e.g., carbon emissions

- Energy: similar to just-in-time SC, with no inventory and instantaneous delivery of goods; smart grid with focus on storage (inventory), renewable energy, e.g, wind solar (supply uncertainty) and real time pricing with smart device (pricing).

- Healthcare: flows of people, expertise, money resources; coordination of conflicting objectives: hospitals, doctors, insurers, pharmaceutical, device companies, patients.

For the later sections, we will study all the elements from downstream to upstream one by one.

2 Customer Demand Modeling

2.1 Bass Diffusion

We first investigate the demand from customers in the aggregation level. Consider the lift cycle of a product, we'd like to analyze its sales along time and thus understand its market implications. 2 shows the variation of sales in four stages, which is pretty much close to the shape of virus contagion.

Bass diffusion curve [1] decomposes new adopters of product into two components: innovators and imitators, which are two different customer segments. It assumes that over the lifetime of the product, total m purchases will be made, and along the time, the probability of new adoption is a linear function of the previous buyers

$$P(t) = p + q\frac{D(t)}{m}$$

where $P(t)$ = Prob(purchase at $t|$ not purchased yet by t), p = coefficient of innovation, q = coefficient of imitation, m = market size and D(t) = Number of people who have made a purchase by t. Usually we have $p << q$ but it can be violated. Then we can derive the result as follows.

Define the demand rate at t:

$$d(t) = \frac{\partial D(t)}{dt}$$

Then

$$P(t) = \frac{d(t)}{m - D(t)} \Rightarrow d(t) = \left(p + \frac{q}{m}D(t)\right)(m - D(t))$$

The 4 Life Cycle Stages and their Marketing Implications

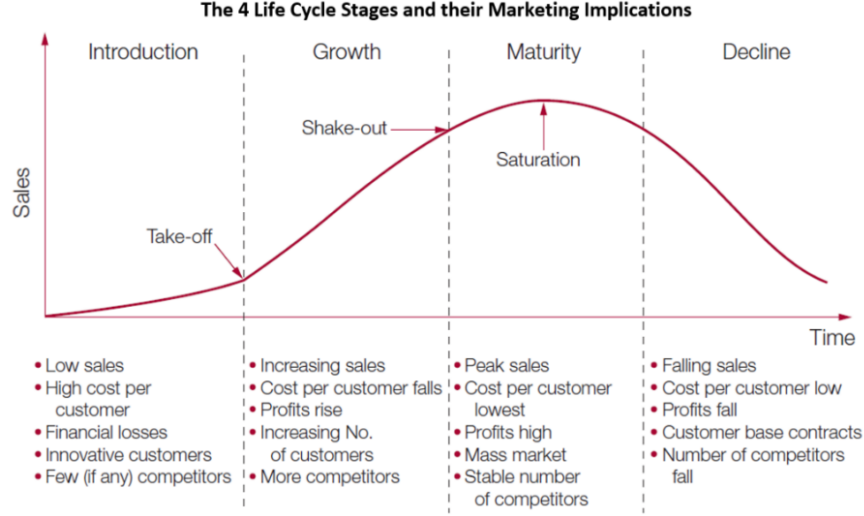| Introduction	Growth	Maturity	Decline
• Low sales	• Increasing sales	• Peak sales	• Falling sales
• High cost per customer	• Cost per customer falls	• Cost per customer lowest	• Cost per customer low
• Financial losses	• Profits rise	• Profits high	• Profits fall
• Innovative customers	• Increasing No. of customers	• Mass market	• Customer base contracts
• Few (if any) competitors	• More competitors	• Stable number of competitors	• Number of competitors fall

Figure 2: Product Life Cycle

By solving differential equation, we get closed-form solution

$$D(t) = m\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}, \quad d(t) = m\frac{p(p+q)^2 e^{-(p+q)t}}{\left(p + qe^{-(p+q)t}\right)^2}$$

Notice that the expression is linear in m, we can alternatively present the result without the presence of m

$$\frac{f(t)}{1 - F(t)} = P(t) = p + \frac{q}{m}D(t) = p + qF(t), D(t) = mF(t), F(t) = \int_{\tau=0}^{t} f(\tau)d\tau$$

By optimality condition, we can take the derivative and the peak demand(maximal demand rate) occurs at time

$$t^* = \frac{1}{p+q}\ln\left(\frac{q}{p}\right)$$

corresponding demand rate and cumulative demand at t^* are

$$d(t^*) = \frac{m(p+q)^2}{4q}, \quad D(t^*) = \frac{m(q-p)}{2q}$$

An important observation is about the relation between p and q. Normally as we mentioned when $p < q$, there is a standard demand curve with some optimal time but when $q < p$, the demand curve is decreasing all the time, which means the imitation effect is negligible and the product is somehow unsuccessful.

After understanding the structural property of Bass model, we still need to estimate its parameters given some historical data. By discretizing the model, we can get the expression

$$d(t) = pm + (q - p)D(t) - \frac{q}{m}D(t)^2$$

which is a linear form if we substitute the coefficients and recover original variables after solving linear regression.

2.2 Strategic Consumer Behavior

From the Bass model we understand the life curve of a new product without considering the effect of price. Normally, we will see the phenomenon that prices of new products decline in order to attract more consumer segments, which is referred as price skimming. Actually, there are some other factors leading to price decline, such as newer version of product, entry of competitor, clearance etc, but we focus on price skimming to better understand the interplay between seller and consumer.

Let us consider the model with settings in [2]

- A monopolist seller of durable new products has to set prices in T periods with constant marginal production cost c.

- N heterogeneous consumers will potentially purchase at most 1 unit of product and they have random valuation from $U(0, v^+)$.

- Both seller and consumers share a common discount factor $\delta \in (0, 1)$.

The game between seller and consumers plays as follows: in period t, seller observes the market state of last period v_{t-1}, which denotes the lowest valuation of some critical consumer who has purchased before t, and then he will set a price p_t; consumers then observe the price p_t, along with rational expectations for pricing policy of seller, and they decide whether to purchase the product. In this way, the market state v_t will be updated and a new round follows. Essentially, we are interested to figure out the Subgame Perfect Nash Equilibrium under this circumstance.

Before analyzing the game, we have to clarify some notations:

- $p_t^* (v_{t-1}) = $ monopolist's equilibrium pricing strategy when faced with state v_{t-1} in period t.

- $v_t^* (p_t, v_{t-1}) = $ lowest reservation value that would purchase in period t, when the monopolist announces price p_t in state v_{t-1}.

- $v_t^* (v_{t-1}) = v_t^* (p_t^* (v_{t-1}), v_{t-1})$.

On the one hand, for any $t \in \{1, \ldots, T\}$ for any $v_{t-1} \in [0, v^+]$, $p_t^* (v_{t-1})$ constitutes the optimal solution for seller in the optimization problem

$$\max_{p_b \ldots, p_T} \sum_{s=t}^{T} \delta^{s-t} (p_s - c) (v_{s-1} - v_s) \frac{N}{v^+}$$

subject to

$$v_t = v_t^* (p_t, v_{t-1})$$
$$v_s \leq v_{s-1}, \quad s = t, \ldots, T$$
$$v_{s+1} = v_{s+1}^* (v_s), \quad s = t, \ldots, T$$
$$p_{s+1} = p_{s+1}^* (v_s), \quad s = t, \ldots, T$$

the objective function maximizes profits over remaining time periods with discount; the first constraint supposes consumers will response rationally at t for any given price; the second constraint preserves monotonicity; the last two constraints just capture the rationality for both seller and consumers in remaining periods.

On the other hand, for any $t \in \{1, \ldots, T\}, v_{t-1} \in [0, v^+], p_t \geq 0$, and $v \in [0, v_{t-1})$, a consumer with reservation value v purchases in period t if and only if his period t utility exceeds his utility from purchasing in periods $t+1, \ldots, T$, or not purchasing at all:

$$
\begin{aligned}
v - p_t &\geq \delta \left[v - p_{t+1}^* \left(v_t^* \left(p_t, v_{t-1} \right) \right) \right] \\
v - p_t &\geq \delta^2 \left[v - p_{t+2}^* \left(v_{t+1}^* \left(v_t^* \left(p_t, v_{t-1} \right) \right) \right) \right] \\
v - p_t &\geq \delta^{T-i} \left[v - p_T^* \left(v_{T-1}^* \left(\cdots \left(v_{t+1}^* \left(v_t^* \left(p_t, v_{t-1} \right) \right) \right) \cdots \right) \right) \right] \\
v - p_t &\geq 0
\end{aligned}
$$

In addition, we must have $v_t^* \left(p_t, v_{t-1} \right) = \inf \left\{ v \in [0, v_{t-1}) : v \text{ satisfies constraints above} \right\}$.

To characterize the SPNE, we get through three steps: first guess the form of marginal/critical consumer at time t, then solve the seller's optimization problem as DP by induction, and finally verify the DP solution is consistent with consumers' rationality. After getting the optimal pricing strategy for the seller, we can directly show that under the SPNE, the optimal price monotonically decrease: $p_{t+1}^* < p_t^*$, which verifies the price skimming effect. Moreover, the optimal price at each time period should be less than myopic solution. This result is quite intuitive because the seller knows that consumers are rational, he should propose lower price than static solution. On top of that, it can shown that with long enough horizon T, the SPNE solution will converge in some way.

3 Facility Location Modeling

Once having different kinds of models to characterize customer demands, we need to construct the facilities that serve the customer(fulfill the demand), which leads to facility location problem. Normally, the high-level consideration in this topic lies in the trade-off between facility cost and transportation cost(serving cost). Therefore, we mainly care about where to build a facility and which customer to serve. Some common notations and parameters are shown as follows:

- $h_i :=$ demand size at location i
- $d_{ij} :=$ distance between location i and j
- $f_j :=$ facility cost at location j
- $a_{ij} :=$ whether a potential facility at location i can cover customer at location j
- $x_i :=$ whether facility is build at location i
- $y_{ij} :=$ whether facility i should serve location j
- $z_i :=$ whether customer location i is served

3.1 P-Median

The P-Median model aims to minimize the demand weighted transportation cost over all customers for p facilities.

$$\min \sum_i \sum_j h_i d_{ij} y_{ij}$$

$$\text{s.t.} \sum_j x_j = p$$

$$\sum_j y_{ij} \geq 1 \quad \forall i$$

$$y_{ij} \leq x_j \quad \forall i, j$$

$$\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^n$$

Remark. In this model, $\boldsymbol{y}$ can be relaxed to be continuous.

3.2 Set Covering

Instead of transportation cost, the Set Covering model aims to minimize the facility cost.

$$\min \sum_j f_j x_j$$

$$\text{s.t.} \sum_j a_{ij} x_j \geq 1 \quad \forall i$$

$$\boldsymbol{x} \in \{0, 1\}^n$$

3.3 P-Covering

The P-Covering model aims to maximize the coverage of demands for p facilities.

$$\max \sum_i h_i z_i$$

$$\text{s.t.} \sum_j x_j = p$$

$$\sum_j a_{ij} x_j \geq z_i \quad \forall i$$

$$\boldsymbol{x}, \boldsymbol{z} \in \{0, 1\}^n$$

3.4 P-Center

The P-Center model aims to minimize the maximal distance between any pair of facility and customer, such as some public service design(emergency).

$$\min \quad w$$
$$\text{s.t.} \sum_j x_j = p$$
$$\sum_j y_{ij} = 1 \quad \forall i$$
$$y_{ij} \le x_j \quad \forall i, j$$
$$\sum_j d_{ij} y_{ij} \le w \quad \forall i$$
$$\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^n$$

Remark. In this model, $\boldsymbol{y}$ can be relaxed to be continuous.

3.5   P-Dispersion

Sometimes facilities may interfere with each other so that customer service quality is affected. The P-Dispersion model aims to maximize the minimal distance between between facilities.

$$\max \quad w$$
$$\text{s.t.} \sum_j x_j = p$$
$$d_{ij} + M(2 - x_i - x_j) \ge w \quad \forall i, j$$
$$\boldsymbol{x} \in \{0, 1\}^n$$

3.6   P-Maxisum

Actually, not all the facilities are built to literally serve the customer demand such as hazardous facility. Instead, p facilities have to be build within some potential locations but we need to maximize the demand weighted distance between the customers and nearest facility.

$$\max \quad \sum_i \sum_j h_i d_{ij} y_{ij}$$
$$\text{s.t.} \sum_j x_j = p$$
$$\sum_j y_{ij} = 1 \quad \forall i$$
$$y_{ij} \le x_j \quad \forall i, j$$
$$\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^n$$

Till now, the formulation is quite similar to P-Center, but unfortunately, since we are doing maximization, the nearest facility requirement may not hold directly. There are two ways of adding constraints for this concern

- $d_{ij}Y_{ij} \leq d_{ik}X_k + M\left(1 - X_k\right), M$ is a large number, $\forall i \in I, j \in J, k = 1, \cdots, |J|$

- $\sum_{k=1}^{m} Y_{i[k]_i} \geq X_{[m]_i}, \forall i \in I, m = 1, \cdots, |J|$

3.7   Hub and Spoke

For the models above, demand is characterized at each customer node while in practice, we may have more complex situation such as demand of origin-destination pair in transportation, especially airline industry. A common transportation network called Hub and Spoke aims to minimize the total transportation cost by serving all the demands of OD pairs. In particular, each customer node has to be assigned to exactly one hub node and the cost between hub nodes has discount because of economies of scale.

$$
\begin{aligned}
\max \quad & \sum_i \sum_j h_{ij}(\sum_k d_{ik}y_{ik} + \sum_m d_{mj}y_{jm}) + \alpha \sum_k \sum_m d_{km}y_{ik}y_{mj}) \\
\text{s.t.} \quad & \sum_j x_j = p \\
& \sum_j y_{ij} = 1 \quad \forall i \\
& y_{ij} \leq x_j \quad \forall i, j \\
& \boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n
\end{aligned}
$$

The constraints are the same as problems before but we adjust objective to capture the cost of each route, which has to be linearized.

3.8   Competition Set Covering

Suppose when setting up the facilities, some nodes have been covered by competitors so we aim to maximize the market share by locating p stores. In this case, we have to define some new parameters:

- $b_{ij} :=$ whether the distance between customer i and node j is strictly less than its nearest competitor's store

- $c_{ij} :=$ whether the distance between customer i and node j is less than its nearest competitor's store

The reason we need these two different parameters comes for the situation that the new store is built at the same location as competitor's and the market share will be divided.

$$\min \sum_i h_i(z_i + 0.5v_i)$$

$$\text{s.t.} z_i + v_i \leq 1 \quad \forall i$$

$$\sum_j b_{ij}x_j \geq z_i \quad \forall i$$

$$\sum_j c_{ij}x_j \geq v_i \quad \forall i$$

$$\sum_j x_j = p$$

$$\boldsymbol{z}, \boldsymbol{v} \in \{0,1\}^n$$

3.9   Uncapacitated Facility Location Problem

$$\min \sum_j f_j x_j + \sum_i \sum_j h_i c_{ij} y_{ij}$$

$$\text{s.t.} \sum_j y_{ij} = 1 \quad \forall i$$

$$x_j \geq y_{ij} \quad \forall i, j$$

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

$$x_j \in \{0,1\} \quad \forall j$$

It's hard to solve it directly and we can come up with some iterative algorithm based on Lagrangian relaxation. Specifically, we first calculate a lower bound $Z_{LR}(\lambda)$ of Z^* with some λ; then we construct a feasible solution based on the relaxed solution, which performs an upper bound; finally we update λ by a subgradient step to reach next iteration.

$$Z_{LR}(\lambda) = \min \sum_j (f_j x_j + \sum_i (h_i c_{ij} - \lambda_i) y_{ij}) + \sum_i \lambda_i$$

$$\text{s.t.} \ x_j \geq y_{ij} \quad \forall i, j$$

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

$$x_j \in \{0,1\} \quad \forall j$$

Obviously the problem can be decomposed by j and the solution of subproblem is easy to get by reasoning. Since $\boldsymbol{x}^*(\lambda)$ is always feasible for the original problem but not the case for $\boldsymbol{y}^*(\lambda)$, we can adjust all the customers to visit its nearest facility and thus the solution becomes feasible and provides an upper bound. Next, for each λ_i, it is easy to verify the subgradient is $1 - \sum_j y_{ij}$, and we can derive the following update rule

$$\lambda_i^{t+1} = \lambda_i^t + \Delta^t (1 - \sum_j y_{ij})$$

where $\Delta^t = \alpha^t (UB - LB)/\sum_i (1 - \sum_j y_{ij})^2$.

4 Demand Uncertainty

Even though we are able to characterize customer demand from many perspectives, it's still crucial to deal with demand uncertainty for better operational decisions. There are two basic notions about demand uncertainty: service level and demand variance. Service level α describes the probability that demand can be completely fulfilled, which is highly affected by demand variance σ^2. Essentially, when customer demand has high variance, supplier need to hold more safety stock correspondingly to guarantee some service level.

4.1 Risk Pooling

Consider a standard problem that there are multiple types of demand(customers), with different normal distributions respectively, and we need to decide the mechanism of serving these demands. The most naive design is decentralized system that each type of customer is served by one exclusive supplier, and a typical alternative is centralized system that all types of customers are pooled together by a central supplier. Our objective is to analyze the cost of these two systems and understand the benefit of risk pooling.

Let's consider the model with setting

- Supplier uses base stock policy for inventory operations

- Holding cost is h and penalty cost is p for each supplier

- Demands can be correlated so that $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$

We need to decide the base stock policy S_i in the infinite horizon.

For the decentralized system, the expected cost for supplier at each period is

$$g(S_i) = h\mathbb{E}\left[(S_i - x)^+\right] + p\mathbb{E}\left[(x - S_i)^+\right]$$

It's easy to take the derivative and get optimal solution by switching integral and derivative

$$g'(S_i) = h\mathbb{E}\left[\mathbf{1}(S_i > x)\right] + p\mathbb{E}\left[-\mathbf{1}(S_i < x)\right] = hF(S_i) - p(1 - F(S_i))$$

which leads to $S_i^* = F_i^{-1}(\frac{p}{p+h}) = \mu_i + Z_\alpha \sigma_i$ and that $g(S_i^*) = (p+h)\sigma_i \phi(Z_\alpha) = \eta \sigma_i$. Finally we sum up all suppliers to get total cost

$$G(D) = \eta \sum_i \sigma_i$$

For the centralized system, we can aggregate the demand as a new normal distribution $\mathcal{N}(\sum_i \mu_i, \sum_i \sum_j \sigma_{ij})$ so that the optimal base stock level is $S = \sum_i \mu_i + Z_\alpha \sqrt{\sum_i \sum_j \sigma_{ij}}$, and total cost is

$$G(C) = \eta \sqrt{\sum_i \sum_j \sigma_{ij}}$$

Theorem 4.1. $G(C) \leq G(D)$

In particular, we can discuss three scenarios: first, if demands are positively correlated, the benefit of pooling is minor especially completely correlated; second, if demands are not correlated, centralized system helps; third, if demands are negatively correlated, pooling effect is obvious but safety stock can never be eliminated.

4.2 Postponement

Let's extend the problem to the case when lead time is necessary for manufacturing. In this case, when dealing with multiple types of demands, the supplier can first manufacture a generic product for t periods, and then diversify to different end products for remaining T periods. For ease of exposition, we assume the demands of end products are independent and holding cost for end products are always larger than that of generic product. We need to decide the optimal timing t that switch from generic to end product.

Similar with previous model, in order to achieve service level α, the safety stock of the generic product should be $Z_\alpha \sqrt{t \sum_i \sigma_i^2}$ while that of the end product i should be $Z_\alpha \sqrt{(T-t)\sigma_i^2}$. The expected cost in terms of t can be calculated as

$$G(t) = h_0 Z_\alpha \sqrt{t \sum_i \sigma_i^2} + \sum_i h_i Z_\alpha \sqrt{(T-t)\sigma_i^2}$$

which leads to the optimal timing $t^* = T$. Intuitively, it means we should always tend to aggregate the demands unless they have to be diversified.

4.3 Transshipment

In many practical setting, many other costs should be taken into consideration for the pooling system, such as transportation cost, ordering cost, etc. Let's consider a decentralized system with two suppliers and two customers, but the transshipment between suppliers is allowed to balance the inventory if necessary. Retailer i holds a base-stock policy S_i with holding cost h_i and penalty cost p_i. The ordering cost from retailer i is c_i while the transshipment cost from i to j is c_{ij}.

Several assumptions are needed to prevent trivial cases by defining the transshipment surplus $\delta_{ij} = c_{ij} + c_i - c_j$

- $\delta_{ij} > 0$

- $\delta_{ij} \le p_j + h_i$

- $\delta_{ij} \ge p_j - p_i$

- $\delta_{ij} \ge h_i - h_j$

Then the expected cost can be represented as

$$g(S_1, S_2) = \sum_i (c_i \mathbb{E}[Q_i] + \sum_{j \ne i} \mathbb{E}[Y_{ij}] c_{ij} + h_i \mathbb{E}[IL_i^+] + p_i \mathbb{E}[IL_i^+])$$

We have to transform all random variables in the expression of S_i by discussing several scenarios. For example, when dealing with Y_{ij}, we have to consider

- $d_i \ge S_i$

- $d_i < S_i, S_i + S_j - d_i \ge d_j > S_j$

- $d_i < S_i, S_i + S_j - d_i < d_j$

Theorem 4.2. $\alpha_i(S_1, S_2) = \alpha_i^0(S_1, S_2) + |\frac{d\mathbb{E}[Y_{ji}]}{dS_i}|$

References

[1] Frank M Bass. "A new product growth for model consumer durables". In: *Management science* 15.5 (1969), pp. 215–227.

[2] David Besanko and Wayne L Winston. "Optimal price skimming by a monopolist facing rational consumers". In: *Management science* 36.5 (1990), pp. 555–567.