Optimal Buying & Selling Decisions

Howard Hao-Chun Chuang (莊皓鈞)

Professor
College of Commerce
National Chengchi University

November, 2024
Taipei, Taiwan

1. 今天要拿多少報紙來賣

The Newsvendor Problem

Q (Quantity) under Random Demand D.

- **Optimal prediction ≠ optimal decision** 5個 level，100天

an item has random demand D = 1, 2, 3, 4, or 5 with prob=0.2 each

unit cost c=25; unit price p=100

for the next 100 days, every morning we stock Q units when we open the store and discard(報廢) anything that is unsold by the end of a day

隨機

optimal order quantity (Q^*) under stochastic demand (D)?

缺貨成本

準備的量小於需求時，代價是多少

what is the unit cost for $Q<D$ (under-stock)? $c_u = p$(價格) – c(成本) = 75

準備的量大於需求時，代價是多少

what is the unit cost for $Q>D$ (over-stock)? $c_o = c$(成本) – s(殘值) = 25

過量庫存成本

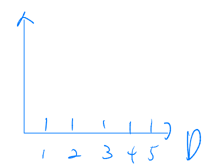
$$c_o * P(D \le Q^*) \ge c_u * P(D > Q^*)$$

$P(D \le \boxed{Q^*}) = \frac{C_u}{C_u + C_o} = \frac{75}{75+25} = \frac{75}{100} = 0.75$

臨界分數

$c_u/(c_u+c_o)$ is the critical fractile, i.e., optimal service level (服務水準)

xx% service level: (100-xx) glitches out of 100 trials



Q=1 P(D≤1) = 0.2

Q=2 P(D≤2) = 0.4

⋮

Q=4 P(D≤4) = 0.8

有超過臨界值

$C_u \downarrow$

服務水準↓

c_u in practice is oftentimes more than p-c

loss of goodwill, strategic (or even panic) buying, etc.

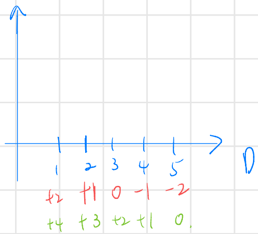
人為上調

$(P-C) * Min(D, Q) + (S-C) * Max(Q-D, 0)$

殘值: Salvage

$$\underset{Q}{Max} \; \underbrace{P * Min(D,Q)}_{Sold} - \underbrace{C * Q}_{Cost} + \underbrace{S * Max(Q-D, 0)}_{Left}$$

殘值：Salvage



Optimal Production = 3

Optimal Production = 5

$$C_o * P(D \le Q) \ge C_u * P(D > Q)$$

$$\Rightarrow P(D > Q) = 1 - P(D \le Q)$$

$$C_o * P(D \le Q) = C_u * [1 - P(D \le Q)]$$

$$(C_o + C_u) * P(D \le Q) = C_u$$

$$P(D \le Q) = \frac{C_u}{C_o + C_u}$$

我要準備多少庫存才能滿足大多數需求？

if $P(D \le Q) = 0.95$

$\Rightarrow$ 最佳訂購量 Q 能滿足 95% 的需求

100 次需求有 95 次被滿足

Optimal prediction

$\neq$

Otimal decision

不能只看需求分布，

要考慮不同情況下的成本

C_u, C_o

* 兩者不相等

$$\frac{C_u}{C_u + C_o}$$

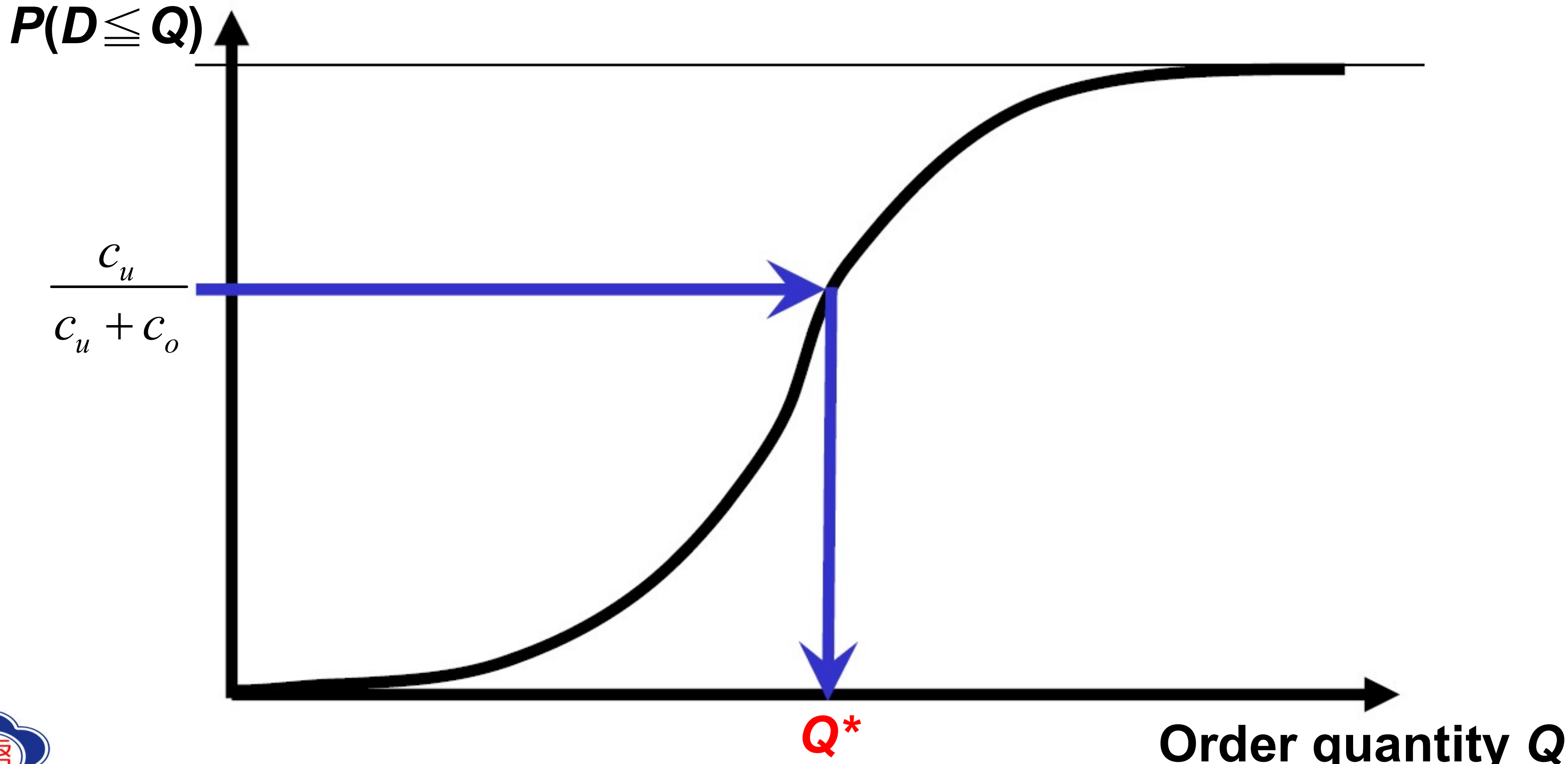
$C_u \uparrow$: 服務水準提高

$C_o \downarrow$: (要準備更多庫存)

$C_o \uparrow$: 服務水準下降

$C_u \downarrow$: (準備之庫存減少)

The Newsvendor Problem



The Newsvendor Problem

$c = 4.2$
$p = 5.7$
$s = 2.4$

- **A fish dealer in the Kanziding fish market**

He buys fresh Taiwan snapper in the midnight for \$4.2 per pound
He sells the fish in the market for \$5.7 per pound
Any remaining fish is sold to a cat food producer for \$2.4 per pound
What is the unit cost of under-ordering?

$C_u = 5.7 - 4.2 = 1.5$

What is the unit cost of over-ordering?

$C_o = \cancel{2.4 - 4.2 = -1.8}$ $4.2 - 2.4 = 1.8$

Daily demand ~ Normal(*mu*=80 pounds, *sigma*=10 pounds)

Where is this distribution from? Is an IID assumption reasonable?

Central limit Theorem

Independent and Identically Distributed (獨立且同分布)

What is the optimal service level?

$\frac{C_u}{C_o + C_u} = \frac{1.5}{1.8 + 1.5} = \frac{1.5}{3.3} = 0.454$

What is the optimal stocking level?

stats.norm.ppf (0.454, $\underbrace{80, 10}_{\text{Normal}}$)

$P(X \le x) = 0.454$

$x = 78.844$

Case Study: Scotia Snowboards

- **The Scotia Corporation will sell NEW snowboards next winter**
  - **Snowboards must be made by September (due to production lead time)**
    - C: **cost of manufacturing is $20 per snowboard**
    - P: **sales revenue is $48 per snowboard**
    - S: **unsold snowboards at the end of the winter have a value of $8 each**

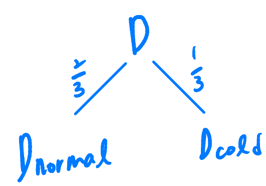
$\begin{cases} C_u = P - C = 28 \\ C_o = C - S = 12 \end{cases} \Rightarrow SC^* = P(D \le Q^*) = \frac{C_u}{C_u + C_o} = 0.7$

  - **Market demand is weather-dependent**
    - ***P*(weather is normal)=2/3; P(weather is cold)=1/3**

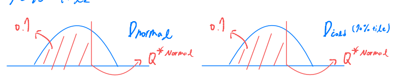
  - **Predictions about demand (*D*) for snowboards, if weather is normal**
    - **probability of 1/4 being below 60,000**
    - **probability of 1/2 being below 75,000**
    - **probability of 3/4 being below 90,000**

  - **Predictions about demand (*D*) for snowboards, if weather is cold**
    - **probability of 1/4 being below 80,000**
    - **probability of 1/2 being below 100,000**
    - **probability of 3/4 being below 125,000**

$\frac{C_u}{C_u + C_o} = \frac{28}{40} = 0.7$

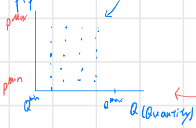


25% tile	60k	25% tile	80k
中位数 = 50% tile	75k	50% tile	100k
75% tile	90k	75% tile	125k



Min → Perf (P, Q)

同時設定 (P, Q)



profit 式子

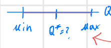
必考

P, V

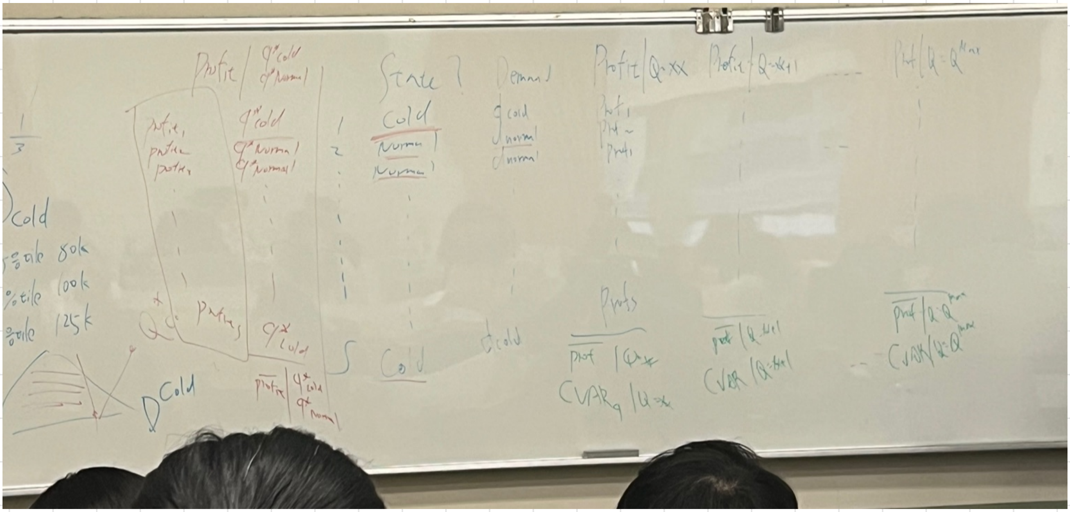
Salvage value

$$\text{Max} \quad \underbrace{p * Min(D, Q)}_{\text{Sold}} - \underbrace{C * Q}_{\text{Cost}} + S * \underbrace{Max(Q - D, 0)}_{\text{Left}}$$

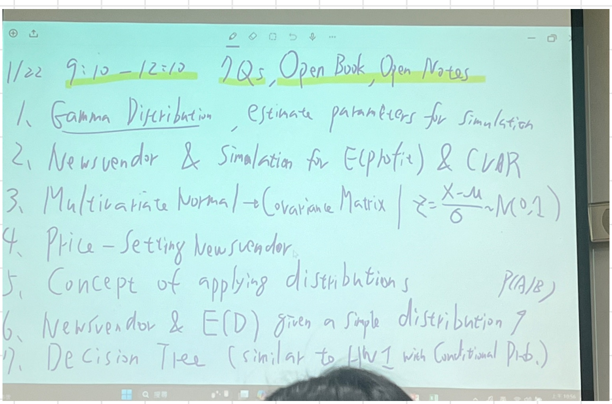
line search



Q candidates : $Q^{small}, \dots, Q^{max}$



1, Gamma Distribution, Estimmate Parameters of simulations
2, Newsvender & Siimulation for E(profile) & CVAR
3. Mulitivariate Normal –> Covariana Matrix | Z = (X – u) / σ ~ N (0,1) 標準化再轉回來
4.Price – Setting Newsvender
5.Concept of applying distributions
6. Newsvendor & E(D) given a sample distribution
7.Decision Tree (similar to HW1 with conditional probability => P(A|B)



Case Study: Scotia Snowboards

Max E (profit)
or
Min CVAR

- The objective is to **maximize expected profits**

  What are the decision variable, parameters, & objective function?

  Q*　　Demand, P, C, S　　Distributions

  D ~ Myerson(q1, q2, q3, lower, upper, tailprob)

  for *d* in [lower, upper]; Thanks goes to Prof. Roger Myerson @ U of Chicago
  q1: xx%tile(xx<50, often xx=25)
  q2: 50%tile(median)
  q3: xx%tile(xx>50, usually xx=75)
  lower: minimum; upper: maximum;
  tailprob: *P*(*D*<q1) + *P*(*D*>q3)

  需求極端區間的机率

  For q1=25%tile, q2=50%tile, & q3=75%tile, the Myerson distribution
  generalizes the normal distribution, if q3-q2=q2-q1 & *d* in [-Inf, Inf]
  generalizes the lognormal distribution, if q3/q2=q2/q1 & *d* in (0, Inf)
  https://stat.ethz.ch/R-manual/R-devel/library/stats/html/Lognormal.html
  https://en.wikipedia.org/wiki/Log-normal_distribution

- 當 q3 − q2 = q2 − q1 且需求 d 在無界區間時，Myerson 分佈退化為正態分佈 (Normal Distribution)。
- 當 q3 / q2 = q2 / q1 且需求 d 在正區間時，Myerson 分佈退化為對數正態分佈 (Log-normal Distribution)。

Case Study: Scotia Snowboards

- **The Scotia Corporation will sell NEW snowboards next winter**
  - **Snowboards must be made by September (due to production lead time)**
    - **cost of manufacturing is \$20 per snowboard**
    - **sales revenue is \$48 per snowboard**
    - **unsold snowboards at the end of the winter have a value of \$8 each**
  - **What are the C_u, C_o, & critical fractile?**
  - **What are the Q^* values for normal and cold weather?**
  - **What if uncertainty about normal or cold weather could be resolved?**
    - **delay the production but pay extra money**
    - **buy weather forecast from the best business intelligence**
  - **What is the value of perfect information about the state of weather?**

知道是左還是右的價值

Q

S D

同一行

同D. RV.

The Newsvendor Problem

- **A single-period model** (Edgeworth 1888, Arrow et al. 1951)
  Applicable for repeated as well as one-shot decisions under uncertainty
    moon cake, down jacket, fresh produce…what else?

    cash reserve of mutual fund for redemption

    seats/rooms in the airline/hospitality industry

    project resources, cloud storage, network bandwidth, etc.

  Big data newsvendor
    with machine learning & predictive analytics

  Small data newsvendor
    with mental accounting & expert speculations

  如何避險不用去做這個隨機最佳化？

例如幫Uniqlo決定冬天衣服的量，越靠近時間在做決策越好，但考能會讓價格提升

1. Delay Q Decisions
2. ↑ $ Find Ways 提高產值
3. Contract (Buy back)
   ⇒ Shift back to vender

背景：Beta 分布參數的公式

Beta 分布的形狀參數 α 和 β 與分布的平均值 μ 和方差 σ^2 的關係如下：

1. 平均值公式：

$$\mu = \frac{\alpha}{\alpha+\beta}$$

2. 方差公式：

$$\sigma^2 = \frac{\alpha \cdot \beta}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)}$$

老師 => $P \times sold - (sold - show) * P * \underset{\text{refound}}{\underline{0.5}} - bump * 2.5 * P$

min(D, Q)

$TP = P \times Min(D, \cancel{Q}) + \underline{RV\_NO(P) \times 0.5 \times P + (1 - RV\_NO(P)) \times P} - 2.5P \times (D - Q)$

Case Study: Starlux Airlines (P, Q)

Price

- Starlux Airlines will launch a Taipei-Naha flight next summer
  - Flight capacity: 150 seats; Fixed cost: \$12,000 USD
  - Demand (D) = *f*(Price) = b0 + b1*Price + noise
    - b0=600 & b1=-5. How are the two parameters estimated?
    - noise ~ Normal(0, 10)

(0, 1)

- Fraction of NO-show is a beta RV (mean=0.1 & stdev. =0.025)
  - each NO-show will receive a refund 50% of Price
  - how do we derive parameters of the beta distribution?
- Cost of overbooking

  $D \times P + \underline{No} \times 0.5P - bump \times 2.5P$

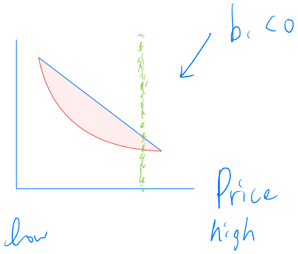
  - each "bumped" customer will have a 250% of Price compensation
  - should Starlux Airlines allow for overbooking?

- K董 asks you to decide on the optimal selling strategy
  - what are the decision variables & objective? Price, 售票數量
  - what are the relevant performance metrics & optimal decisions?

$D = b_0 + b_1 Price + error \sim N(0, 10)$

price elasticity 對價格的敏感度

"空位/超賣率



$P = yy$, $Q = xx$

Demand	Sold	No Show $[0, 1]$	Show up	No show	Dumped
d_1	$Min(Q, d_1)$	0,			
d_2	$Min(Q, d_2)$	0.05	(1 - No show)		
⋮	⋮	⋮			
d_s	$Min(Q, d_s)$	0.08			

Refound $P \times 0.5$...

Penality

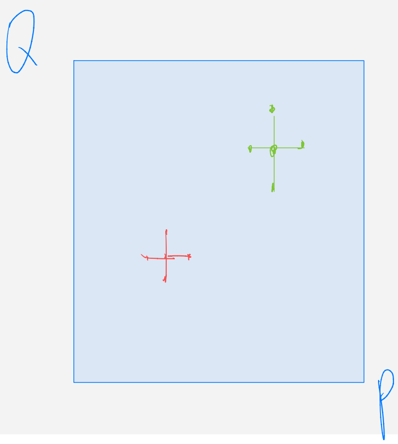
來了沒位置

TP = Revenue + No Show - Over booking - Fixed Cost

Hooke-Jeeves

```
function hooke_jeeves(f, x, α, ϵ, γ=0.5)
    y, n = f(x), length(x)
    while α > ϵ
        improved = false
        x_best, y_best = x, y
        for i in 1 : n
            for sgn in (-1,1)
                x′ = x + sgn*α*basis(i, n)
                y′ = f(x′)
                if y′ < y_best
                    x_best, y_best, improved = x′, y′, true
                end
            end
        end
        x, y = x_best, y_best

        if !improved
            α *= γ
        end
    end
    return x
end
```



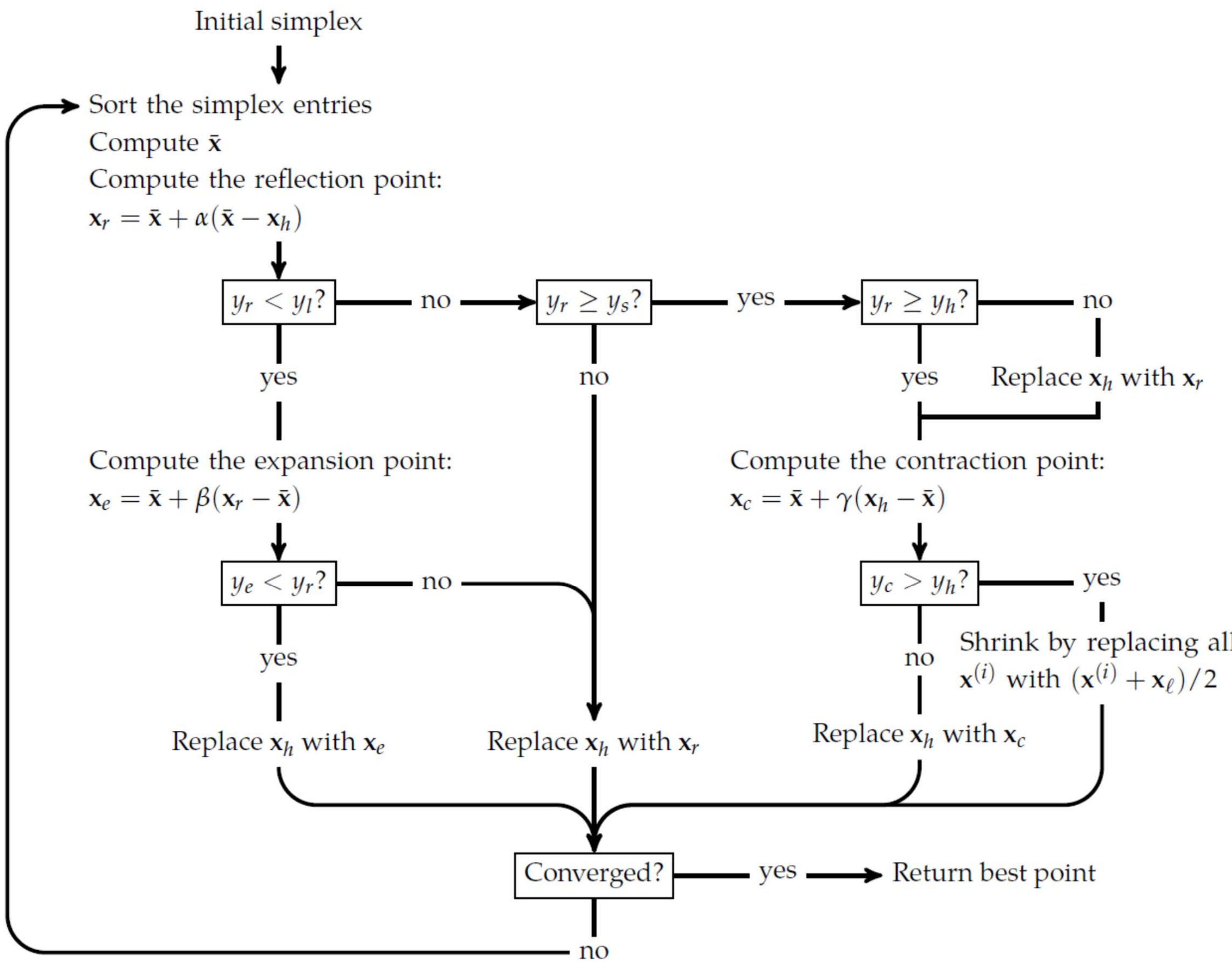
向前,後左,右各算一次

```
basis(i, n) = [k == i ? 1.0 : 0.0 for k in 1 : n]
```

basis(3, 4) = [0, 0, 1, 0]

Algorithm 7.5. The Hooke-Jeeves method, which takes the target function f, a starting point x, a starting step size α, a tolerance ϵ, and a step decay γ. The method runs until the step size is less than ϵ and the points sampled along the coordinate directions do not provide an improvement. Based on the implementation from A. F. Kaupe Jr, "Algorithm 178: Direct Search," *Communications of the ACM*, vol. 6, no. 6, pp. 313–314, 1963.

Nelder-Mead



Initial simplex
Sort the simplex entries
Compute $\bar{\mathbf{x}}$
Compute the reflection point:
$\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_h)$
$y_r < y_l$?
no
yes
$y_r \geq y_s$?
yes
no
$y_r \geq y_h$?
no
yes
Replace $\mathbf{x}_h$ with $\mathbf{x}_r$
Compute the expansion point:
$\mathbf{x}_e = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}})$
Compute the contraction point:
$\mathbf{x}_c = \bar{\mathbf{x}} + \gamma(\mathbf{x}_h - \bar{\mathbf{x}})$
$y_e < y_r$?
no
yes
$y_c > y_h$?
yes
no
Shrink by replacing all $\mathbf{x}^{(i)}$ with $(\mathbf{x}^{(i)} + \mathbf{x}_\ell)/2$
Replace $\mathbf{x}_h$ with $\mathbf{x}_e$
Replace $\mathbf{x}_h$ with $\mathbf{x}_r$
Replace $\mathbf{x}_h$ with $\mathbf{x}_c$
Converged?
yes
Return best point
no

Nelder-Mead

$\min f(x)$ $x_\ell, \ldots, x_3, x_h, \ldots$ 电脑找

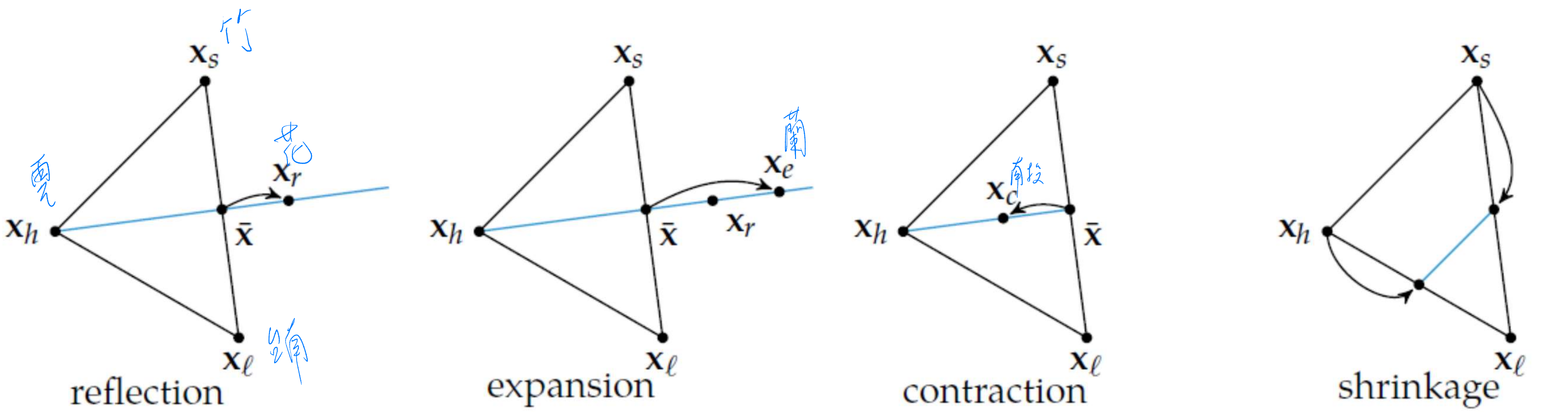
x_ℓ: 最好
x_h: 最壞
x_s: 次壞

Reflection. $\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_h)$, reflects the highest-valued point over the centroid.

Contraction. $\mathbf{x}_c = \bar{\mathbf{x}} + \gamma(\mathbf{x}_h - \bar{\mathbf{x}})$, the simplex is shrunk down by moving away from the worst point.

Expansion. $\mathbf{x}_e = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}})$, like reflection, but the reflected point is sent even further. This is done when the reflected point has an objective function value less than all points in the simplex. Here, $\beta > \max(1, \alpha)$ and is typically set to 2.

Shrinkage. All points are moved toward the best point.



$$\max_{P,Q} \sum_{s=1}^{S} \text{Profit}(P, Q) / S$$

S = simulation

Key Takeaways

- **Three lessons**

  1) The Myerson(q1, q2, q3, lower, upper, tailprob) distribution

  generalizes normal & lognormal distributions

  extremely useful for modeling based on subjective estimates

  2) The buying newsvendor problem

  C_u, C_o, & critical fractile (optimal service level)

  the value of perfect information

  3) The selling newsvendor

  jointly optimize price & quantity decisions

  ***derivative-free* search algorithms**

- **Optimal decision-making in a simulation model**

  simulate payoff-related random variables (e.g., cost, demand)

  compute the payoff for a list of alternative decisions/strategies

  pick the decision/strategy based on your decision criteria

Caveats of Simulation Modeling

- **A Monte-Carlo simulation model**
  - **captures aspects of uncertainty that simple formulas cannot**
    - learn chances & distributions of possible outcomes, not just expected predict values
  - **identifies random variables and create sample values**
    - apply discrete & continuous probability distributions, whatever appropriate
  - **has NO internal optimal decision-making capability**
    - enumerate possible strategies/options, then use the model to evaluate those
  - **does NOT provide precise results due to inherent randomness**
    - use a large number of trails/runs, and construct confidence intervals
- **A successful simulation modeling project depends on**
  - a good understanding of the underlying decision problem
  - one's ability to use the concepts and statistics correctly
  - one's ability to communicate these concepts effectively