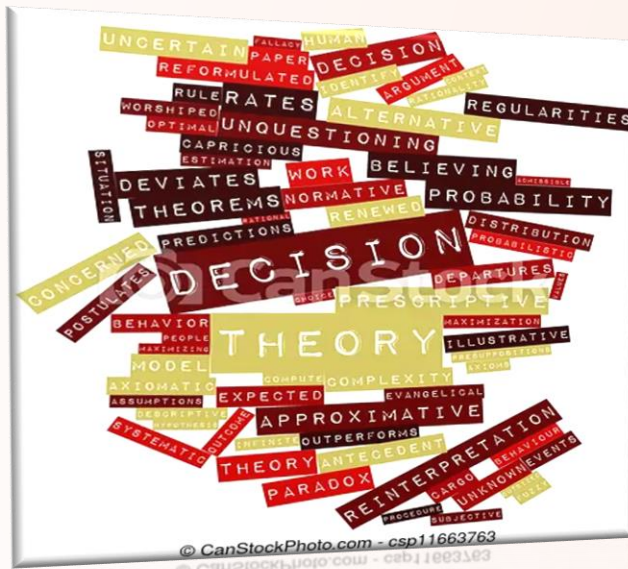


LÝ THUYẾT QUYẾT ĐỊNH

(Decision theory)



Biên soạn: ThS. Phạm Đình Tài
pdtai@ntt.edu.vn
0985.73.39.39

QUYẾT ĐỊNH THIẾU HIỂU BIẾT

- *Sự lợi thế*
- *Nguyên lý Maximin và Leximin*
- *Maximax và luật bi quan-lạc quan*
- *Luật hồi tiếc Minimax*
- *Nguyên tắc không đủ lý do*
- *Hành vi ngẫu nhiên*

1. Sự lợi thế

Trong lý thuyết quyết định, sự thiếu hiểu biết là một thuật ngữ kỹ thuật có ý nghĩa rất chính xác.

- Đề cập đến các trường hợp mà người ra quyết định biết các lựa chọn thay thế là gì và chúng có thể dẫn đến kết quả gì, nhưng không thể gán bất kỳ xác suất nào cho các trạng thái tương ứng với kết quả
- Đôi khi thuật ngữ quyết định không chắc chắn được sử dụng đồng nghĩa.

■ Ví dụ

- Bạn muốn đi du lịch từ Luân Đôn đến Las Vegas. Đường bay này được phục vụ bởi hai hãng hàng không là Black Jack airlines và Air Mojave, cả hai đều có giá vé và mức dịch vụ tương đương nhau.
- Ngay trước khi bạn chuẩn bị đặt vé, bạn đã đọc trên The Times rằng Air Mojave đang trên bờ vực phá sản. Do đó, bạn lập luận như sau: *Không cần biết thông tin về Air Mojave có chính xác hay không, tôi không mất gì khi đặt vé với các hãng hàng không Black Jack. Giá vé và dịch vụ như nhau, nhưng tôi sẽ không phải đối mặt với nguy cơ tiền mất tật mang. Do đó, sẽ hợp lý hơn nếu đặt vé với Black Jack thay vì Mojave.*

■ Ví dụ

- Ma trận quyết định được đưa ra, Bạn nghĩ rằng có một vé hợp lệ sẽ tốt hơn là không có. Mỗi quan hệ thứ tự này có thể được biểu diễn bằng hai số tùy ý, chẳng hạn như 1 và 0.

	Thông tin chính xác	Thông tin không đúng
Các hãng hàng không của Black Jack	Vé hợp lệ	Vé hợp lệ
Air Mojave	Không có vé hợp lệ	Vé hợp lệ



	s_1	s_2
a_1	1	1
a_2	0	1

có thể thấy rõ rằng a_1 chiếm ưu thế hơn a_2

- Để cho là một quan hệ về các hành vi, sao cho $a_i \succeq a_j$ nếu và chỉ khi thực hiện **hành động** i hợp lý hơn là thực hiện **hành vi** j .
- Điều này có nghĩa là bất cứ khi nào $a_i \succeq a_j$ đúng, thì a_i , một cách hợp lý được chọn thay vì a_j .
- Hơn nữa, $a_i \succeq a_j$ có nghĩa là a_i hợp lý như a_j , và $a_i \sim a_j$ có nghĩa là hai hành vi đều hợp lý như nhau.
- Gọi v là một hàm thứ tự gán giá trị cho các kết quả, được coi là các cặp hành vi và trạng thái có thứ tự.
Ví dụ, $v(a_1, s_1) = 1$ có nghĩa là giá trị của việc làm a_1 trong trường hợp s_1 là trạng thái thực của thế giới là 1.

- Nguyên tắc thống trị có hai phiên bản, nguyên tắc thống trị yếu và mạnh

Sự thống trị yếu :	$a_i \succeq a_j$ nếu và chỉ khi $v(a_i, s) \geq v(a_j, s)$ cho mọi trạng thái s
Sự thống trị mạnh mẽ :	$a_i \succeq a_j$ nếu và chỉ khi $v(a_i, s_m) \geq v(a_j, s_m)$ với mọi trạng thái s_m , và có một số trạng thái s_n sao cho $v(a_i, s_n) > v(a_j, s_n)$.

- Nguyên tắc thống trị cũng có thể được áp dụng cho các quyết định có rủi ro.
- Nếu chúng ta biết xác suất của các trạng thái, vẫn có lý khi chọn một phương án thay thế một phương án khác nếu điều đó chắc chắn dẫn đến ít nhất là một kết quả tốt cho dù trạng thái nào trở thành trạng thái thực sự.
- Do đó, không thực sự quan trọng nếu xác suất có liên quan được người ra quyết định biết hay không. Tất cả những gì quan trọng là chúng ta biết rằng ít nhất chúng ta sẽ tốt hơn nếu chúng ta chọn một phương án thay thế khác, bất kể trạng thái nào trở thành trạng thái thực sự.

- Nguyên tắc thống trị cũng có thể được áp dụng cho các quyết định có rủi ro.
- Nếu chúng ta biết xác suất của các trạng thái, vẫn có lý khi chọn một phương án thay thế một phương án khác nếu điều đó chắc chắn dẫn đến ít nhất là một kết quả tốt cho dù trạng thái nào trở thành trạng thái thực sự.
- Do đó, không thực sự quan trọng nếu xác suất có liên quan được người ra quyết định biết hay không. Tất cả những gì quan trọng là chúng ta biết rằng ít nhất chúng ta sẽ tốt hơn nếu chúng ta chọn một phương án thay thế khác, bất kể trạng thái nào trở thành trạng thái thực sự.

2. Nguyên lý Maximin và Leximin



a. Maximin

- Nguyên tắc tối đa tập trung vào kết quả xấu nhất có thể xảy ra của mỗi phương án.
- Theo nguyên tắc này, người ta nên tối đa hóa giá trị tối thiểu có thể đạt được với mỗi hành động.
- Nếu kết quả xấu nhất có thể xảy ra của một phương án tốt hơn phương án khác thì nên chọn phương án trước

- Ví dụ Maximin

	s_1	s_2	s_3	s_4
a_1	6	9	3	<u>0</u>
a_2	<u>-5</u>	7	4	12
a_3	6	4	5	<u>2</u>
a_4	14	<u>-8</u>	5	7

Kết quả xấu nhất có thể xảy ra của mỗi phương án được gạch chân. Tốt nhất trong số các tình huống xấu nhất này đều được gạch chân và **màu xanh**

Quy tắc tối đa quy định rằng người ta nên chọn phương án thay thế là a_3 , bởi vì kết quả xấu nhất có thể xảy ra của hành động đó là **2**, tốt hơn so với các tình huống xấu nhất của tất cả các phương án thay thế khác (0, - 5 và - 8).



- Ví dụ Maximin

Quy tắc tối đa rất dễ chính thức hóa.

Gọi $\min(a_i)$ là giá trị nhỏ nhất có thể đạt được với hành động a_i .

Maximin:

$$a_i \succeq a_j \text{ nếu và chỉ khi } \min(a_i) \geq \min(a_j)$$

b. Leximin

- Nếu các kết quả xấu nhất bằng nhau, người ta nên chọn một phương án thay thế sao cho kết quả tồi tệ thứ hai chắc chắn là tốt nhất có thể.
- Nếu điều này không chỉ ra một hành động duy nhất, thì kết quả tồi tệ thứ ba nên được xem xét, v.v.
- Do đó, trừ khi tất cả các kết quả có thể có của các lựa chọn thay thế hoàn toàn song song, quy tắc tối đa từ vụng (quy tắc leximin) tại một số điểm sẽ chỉ ra một hoặc một số hành vi tốt hơn tất cả các hành vi khác.

Ví dụ Leximin

	s_1	s_2	s_3	s_4
a_1	6	14	8	<u>5</u>
a_2	<u>5</u>	6	100	100
a_3	7	7	7	<u>5</u>

Quy tắc tối đa truyền thống xếp hạng tất cả các lựa chọn thay thế là tốt như nhau, trong khi quy tắc leximin đề xuất điểm a_3 .

Để chính thức hóa quy tắc leximin, hãy đặt $\min^1(a_i)$ là giá trị của kết quả xấu nhất của hành động a_i và $\min^2(a_i)$ là giá trị của kết quả xấu thứ hai của nó và $\min^n(a_i)$ là giá trị của kết quả xấu nhất thứ n của a_i .

Leximin : $a_i \succeq a_j$ nếu và chỉ khi có một số nguyên dương n sao cho

$$\min^n(a_i) > \min^n(a_j) \text{ và } \min^m(a_i) = \min^m(a_j) \text{ với mọi } m < n.$$

- Đối số tốt nhất để chấp nhận tối đa hoặc leximin là gì?
- Có thể lập luận như sau: Theo một nghĩa nào đó, các quy tắc tối đa và leximin cho phép người ra quyết định biến một quyết định thiếu hiểu biết thành một quyết định chắc chắn.
- Chắc chắn, một người ra quyết định chấp nhận quy tắc tối đa hoặc quy tắc leximin không biết kết quả thực sự của lựa chọn của mình sẽ như thế nào.

- Các quy tắc tối đa và leximin quy định rằng a_2 nên được chọn thay vì a_1 .
- Tuy nhiên, từ quan điểm chuẩn tắc, điều này có vẻ hơi kỳ lạ. Hầu hết mọi người chắc chắn sẽ chọn a_1 , bởi vì hầu hết mọi người sẽ đánh giá rằng đáng phải hy sinh 0,1 đơn vị giá trị để có khả năng nhận được 100 đơn vị.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
a_1	1	100	100	100	100	100	100	100	100	100
a_2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1

	s_1	s_1 hoặc s_2 hoặc $s_3 \dots$ hoặc s_{10}
a_1	1	100
a_2	1.1	1.1



3. Maximax và luật bi quan- lạc quan

- Quy tắc tối đa chỉ ra kết quả tồi tệ nhất có thể xảy ra của một quyết định là đặc điểm duy nhất có liên quan về mặt quy chuẩn của nó. Tuy nhiên, tất nhiên người ta cũng có thể chỉ ra kết quả tốt nhất có thể là đặc biệt phù hợp.
- Quy tắc lạc quan - bi quan yêu cầu người ra quyết định xem xét cả kết quả tốt nhất và xấu nhất có thể có của mỗi phương án, sau đó chọn phương án thay thế tùy theo mức độ lạc quan của mình hoặc bi quan.

3. Maximax và luật bi quan- lạc quan

- Giả sử rằng mức độ lạc quan của những người ra quyết định có thể được biểu diễn bằng 1 số thực α từ 0 đến 1, sao cho $\alpha = 1$ tương ứng với mức độ lạc quan tối đa và $\alpha = 0$ là mức độ bi quan tối đa.
- Sau đó, nếu $\max(a_i)$ là kết quả tốt nhất có thể có của phương án a_i và $\min(a_i)$ là kết quả xấu nhất có thể xảy ra, thì giá trị của nó là $\alpha \cdot \max(a_i) + (1 - \alpha) \cdot \min(a_i)$. Đương nhiên, α được giả định là cố định trong suốt quá trình đánh giá tất cả các phương án a_i .

Ví dụ Maximax và luật bi quan- lạc quan

a_1	55	18	28	10	36	100
a_2	50	87	55	90	75	70

- Kết quả tốt nhất có thể có của a_1 là 100 và kết quả xấu nhất có thể là 10. Tức là $\max(a_1) = 100$ và $\min(a_1) = 10$.
- Do đó, nếu người ra quyết định lạc quan ở mức 0,7, thì rằng tổng giá trị của a_1 là $0,7 \cdot 100 + (1 - 0,7) \cdot 10 = 73$. Hơn nữa, kết quả tốt nhất có thể có của phương án a_2 là 90, trong khi kết quả xấu nhất của nó là 50.
- Do đó, giá trị của phương án này là $0,7 \cdot 90 + (1 - 0,7) \cdot 50 = 78$.
- Vì $78 > 73$ nên **phương án a_2 tốt hơn phương án a_1** .

Maximax và luật bi quan- lạc quan

- **lạc quan - bi quan** rõ ràng yêu cầu rằng giá trị có thể được đo lường trên thang đo khoảng thời gian, bởi vì nếu người ta đo giá trị trên thang thứ tự thì sẽ không có ý nghĩa gì khi thực hiện các phép toán số học như nhân và cộng trên các giá trị.
- Tuy nhiên, các quy tắc tối đa và tối đa chỉ yêu cầu giá trị đó được đo trên thang thứ tự.
- Tuyên bố rằng quy tắc lạc quan - bi quan là sự tổng quát hóa của maximin và do đó, tối đa phải được thực hiện với một trên maximin

Maximax và luật bi quan- lạc quan

- Ưu điểm chính của quy tắc lạc quan - bi quan là nó cho phép người ra quyết định đồng thời chú ý đến trường hợp tốt nhất và trường hợp xấu nhất.
- Chỉ số α mô tả tầm quan trọng tương đối của cả hai kết quả. Tất nhiên, sự phản đối rõ ràng là dường như không hợp lý khi hoàn toàn bỏ qua các kết quả ở giữa.

4. Luật hối tiếc Minimax

Ví dụ:

- Bạn được mời đầu tư vào một công ty phần mềm mới ra mắt. Một số người bạn của bạn, những người đã đầu tư vào các công ty tương tự cách đây vài năm, đã kiếm được nhiều tiền. Tất nhiên, bạn có thể mất tiền, nhưng khoản đầu tư khá nhỏ, vì vậy việc công ty có thể phá sản không phải là vấn đề thực sự đáng quan tâm.
- Tuy nhiên, bạn biết rằng nếu bạn **từ chối đầu tư** và công ty trở thành một câu chuyện thành công khác, bạn sẽ cảm thấy rất **hối tiếc**.
- Một số học giả nhận thấy trực giác này quan trọng và cho rằng khái niệm hối tiếc có liên quan đến việc đưa ra quyết định hợp lý. Nói một cách ngắn gọn, những người ủng hộ quy tắc **hối tiếc minimax** cho rằng phương án thay thế tốt nhất là phương án **giảm thiểu sự hối tiếc tối đa**.

Luật hồi tiếc Minimax

Ví dụ:

a_1	12	8	20	20
a_2	10	15	16	8
a_3	30	6	25	14
a_4	20	4	30	10

giả định rằng các kết quả được so sánh trên thang đo khoảng thời gian, không chỉ đơn thuần là thang đo thứ tự. Các số **màu đỏ** biểu thị tương đối thay thế tốt nhất

Luật hồi tiếc Minimax

20 được đánh dấu ở cột ngoài cùng bên phải vì 1 sẽ là lựa chọn thay thế tốt nhất để chọn nếu trạng thái ngoài cùng bên phải xảy ra.

Đối với mỗi kết quả trong ma trận quyết định, các giá trị hồi tiếc có thể được tính bằng cách lấy giá trị của kết quả được đề cập đến trừ đi giá trị của kết quả tốt nhất của mỗi trạng thái (số màu đỏ).

Ví dụ: giá trị hồi tiếc của kết quả thấp nhất trong cột ngoài cùng bên phải là -10, vì $10 - 20 = -10$ và trị hồi tiếc của kết quả thấp nhất thứ hai trong cột ngoài cùng bên phải là -6, vì $14 - 20 = -6$. Bằng cách thực hiện các phép tính tương tự cho tất cả các kết quả trong ma trận quyết định, chúng ta thu được ma trận hồi tiếc:

a_1	12	8	20	20
a_2	10	15	16	8
a_3	30	6	25	14
a_4	20	4	30	10



a_1	-18	-7	-10	0
a_2	-20	0	-14	-12
a_3	0	-9	-5	-6
a_4	-10	-11	0	-10

ma trận hồi tiếc

Luật hồi tiếc Minimax

a_1	-18	-7	-10	0
a_2	-20	0	-14	-12
a_3	0	-9	-5	-6
a_4	-10	-11	0	-10

ma trận hồi tiếc

- Trong ma trận hồi tiếc, các số **màu đỏ** biểu thị giá trị hồi tiếc lớn nhất của mỗi hành động thay thế.
Ví dụ, -18 kém hơn -7, -10 và 0.
- Do đó, -18 là số tiền hồi tiếc cao nhất có được với a_1 .
- Lưu ý rằng tất cả các giá trị hồi tiếc sẽ luôn âm hoặc bằng không. ***Không có cái gọi là hồi tiếc tích cực.***

5. Nguyên tắc không đủ lý do

- Một số nhà tư tưởng lỗi lạc, bao gồm Bernoulli (1654 - 1705) và Laplace (1749 - 1827), đã bảo vệ một quy tắc được gọi là nguyên tắc không đủ lý trí .
- **Ví dụ:** bạn đã đặt một chuyến du lịch cuối tuần đến một đất nước xa lạ. Thật không may, bạn không biết gì về thời tiết tại điểm đến của bạn. Những loại quần áo bạn nên đi? Trong tủ quần áo của bạn, bạn có quần áo cho ba loại thời tiết. Như: thời tiết nóng, vừa và lạnh.
- Ma trận quyết định được đưa ra:

	Nhiệt độ thực là 100 ° F	Nhiệt độ thực là 50 ° F	Nhiệt độ thực là 0 ° F
Mang theo quần áo phù hợp với 100 ° F	15	0	- 30
Mang theo quần áo phù hợp với nhiệt độ 50 ° F	0	15	0
Mang theo quần áo phù hợp với 0 ° F	- 15	0	15

Nguyên tắc không đủ lý do

- Nguyên tắc không đủ lý do quy định rằng nếu một người không có lý do gì để nghĩ rằng một trạng thái của thế giới có khả năng xảy ra cao hơn trạng thái khác, thì tất cả các trạng thái phải được gán xác suất bằng nhau .
- Ví dụ trên có ba trạng thái có thể xảy ra, do đó xác suất được gán cho mỗi trạng thái sẽ là $1/3$. Nói một cách tổng quát hơn, nếu có n trạng thái thì xác suất p được gán cho mỗi trạng thái sẽ là $1/n$.

	$S_1 (1/3)$	$S_2 (1/3)$	$S_3 (1/3)$
a_1	15	0	- 30
a_2	0	15	0
a_3	- 15	0	15

Nguyên tắc không đủ lý do

	$S_1 (1/3)$	$S_2 (1/3)$	$S_3 (1/3)$
a_1	15	0	- 30
a_2	0	15	0
a_3	- 15	0	15

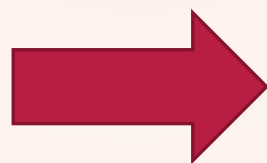
- Giá trị mong đợi của việc mang quần áo phù hợp ở 100°F là $1/3 \cdot 15 + 1/3 \cdot 0 + 1/3 \cdot -30 = -5$.
- Giá trị mong đợi của giải pháp thay thế thứ hai, tức là mang quần áo phù hợp đối với 50°F , là $1/3 \cdot 0 + 1/3 \cdot 15 + 1/3 \cdot 0 = 5$
- Giá trị dự kiến của phương án thứ ba là $1/3 \cdot -15 + 1/3 \cdot 0 + 1/3 \cdot 15 = 0$.

→ Do đó, chọn giải pháp thay thế thứ hai là hợp lý, tức là mang quần áo thích hợp ở 50°F .

6. Hành vi ngẫu nhiên

- Thay vì chọn một hành động hơn một hành động khác, thay vào đó, một người có thể tung đồng xu, ví dụ: chọn phương án thay thế đầu tiên nếu nó hướng lên và hành động thứ hai nếu không có hướng lên.
- Ngẫu nhiên giữa a_1 và a_2 có nghĩa là một giải pháp thay thế mới được đưa vào, đó là hành động ngẫu nhiên giữa các hành vi khác, do đó chuyển ma trận quyết định ban đầu thành một ma trận mới với ba lựa chọn thay thế.
- Nếu ngẫu nhiên được thực hiện bằng cách tung một đồng xu công bằng, ma trận quyết định mới sẽ là:

a_1	1	0
a_2	0	1



a_1	1	0
a_2	0	1
a_3	1/2	1/2

6. Hành vi ngẫu nhiên

- Trong trường hợp bạn biết rằng sẽ phải đối mặt với cùng một loại quyết định hết lần này đến lần khác.
- Do đó, nếu chọn ngẫu nhiên giữa a_1 và a_2 , trung bình sẽ nhận được $1/2$.
- Tuy nhiên, Nếu một người phải đối mặt với cùng một vấn đề quyết định lặp đi lặp lại, theo nghĩa đen, thì sự thật về hai trạng thái sẽ được khắc phục một lần và mãi mãi.
- Vì vậy, sẽ tốt hơn nếu chọn a_1 hoặc a_2 ngay lần đầu tiên và sau đó điều chỉnh hành vi của những hành vi đó trong tương lai theo kết quả của quyết định đầu tiên.
- *Nếu một người chọn a_2 và nhận được 0, người ta nên chuyển sang a_1 , điều này chắc chắn sẽ dẫn đến a_1 trong tất cả các quyết định còn lại.*

THANK YOU !

