博弈论

第二讲：二人零和博弈

黄嘉平

深圳大学中国经济特区研究中心

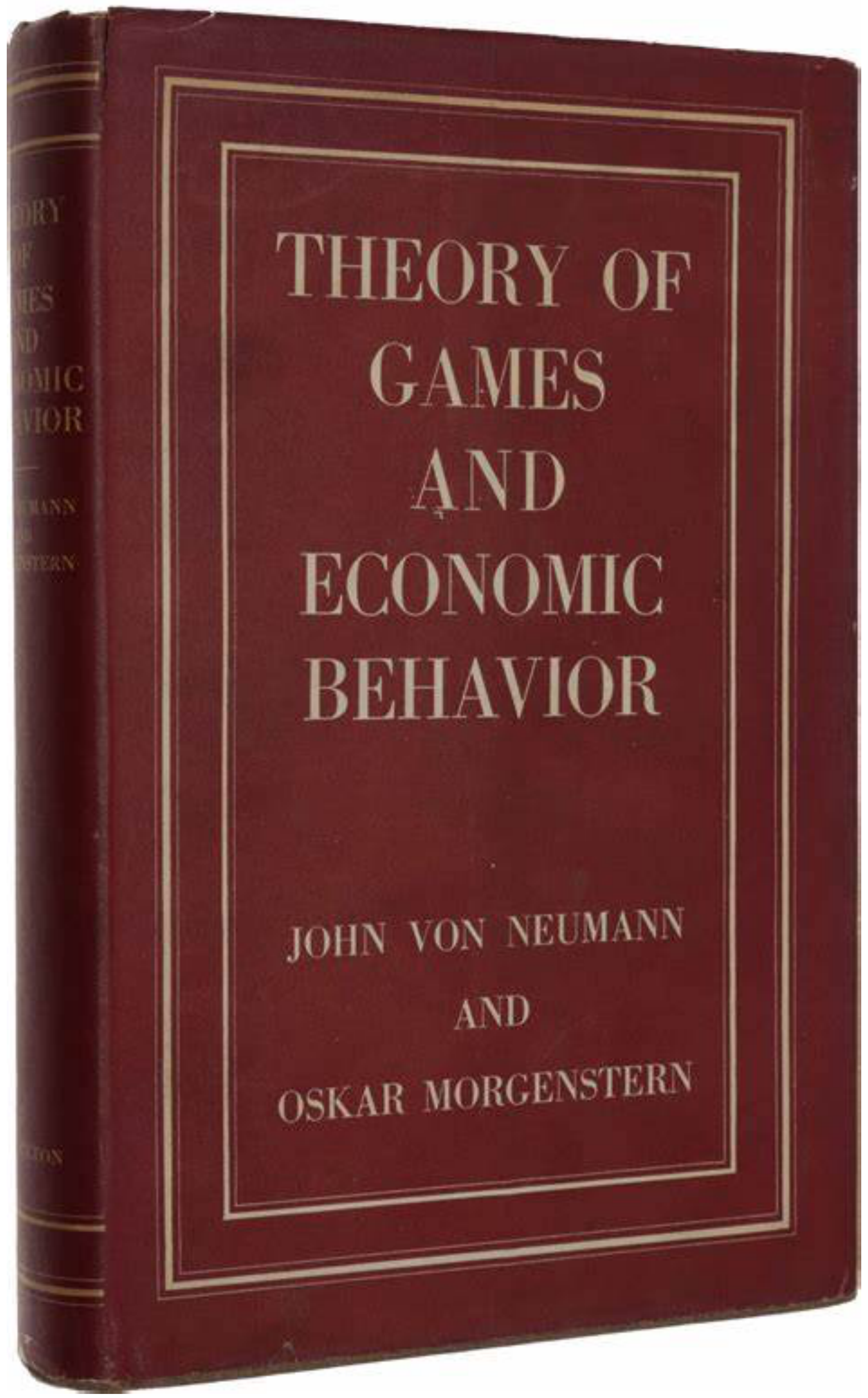
办公地点： 粤海校区汇文楼 1510
丽湖校区守正楼 A 座 3 楼公共办公室

电子邮箱： huangjp@szu

课程主页： https://huangjp.com/GT/

最古老的博弈类型：二人零和博弈

- 零和博弈指参与人的收益之和永远为零的博弈。二人零和博弈的实例包括所有二人棋牌类游戏和两人之间的打赌。
- 最早的关于博弈论的论文被认为是 **Ernst Zermelo** (1913) Über eine Anwendung der Mengenlehre auf die Theorie des Schachpiels (On an Application of Set Theory to the Theory of the Game of Chess)。这是一篇用数学方法分析国际象棋的论文。
- 学术界广泛认为 **John von Neumann** (1928) Zur Theorie der Gesellschaftsspiele (The Theory of Games of Strategy) 是现代博弈论的起源，它证明了二人零和博弈中的 minimax 定理。实际上它并不是最早的，只是其他更早出现的文章被埋没在历史中。
- **John von Neumann** & **Oskar Morgenstern** (1944) *Theory of Games and Economic Behavior* 将 von Neumann (1928) 的内容进一步扩展。这本书真正掀起了博弈论研究的热潮。书中用了大部份篇幅分析零和博弈，只用了一小部讨论合作博弈。



俾斯麦海海战博弈

- 俾斯麦海海战博弈就是一个二人零和博弈：
  有两个参与人，其中一人的收益就是另一人的损失。

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

- 参与人都是理性的，都将最大化自身收益作为行为准则。但是双方的收益是此消彼长，因此自身收益的增加必然对应着对方收益的减少。
- 在这种情况下，一种保守但合理的决策方式是：假设对方会尽可能的减少我方的收益（因为对方要增加自己的收益），我方则在此基础上选择收益最大的行动。例如，盟军可以做出下列分析：
  - 如果我方选择北线：日方选择北线或者南线，我方收益为 2
  - 如果我方选择南线：日方选择北线，我方收益为 1
  - 因此，我方应当选择收益大的，即北线。

基本定义

矩阵博弈
Matrix game

		日军 北线	日军 南线
盟军	北线	2	2
盟军	南线	1	3

俾斯麦海海战博弈

- **矩阵博弈**是一个 $m \times n$ 的实矩阵 A，其中 m 是行数，n 是列数。
- 参与人 1 的（混合）**策略**是关于行的概率分布 $\boldsymbol{p}$，即下面集合的一个要素

$$\Delta^m := \left\{ \boldsymbol{p} = (p_1, \ldots, p_m) \in \mathbb{R}^m \;\middle|\; \textstyle\sum_{i=1}^m p_i = 1,\ p_i \geq 0 \text{ for all } i = 1, \ldots, m \right\}$$

- 参与人 2 的（混合）**策略**是关于列的概率分布 $\boldsymbol{q}$，即下面集合的一个要素

$$\Delta^n := \left\{ \boldsymbol{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n \;\middle|\; \textstyle\sum_{j=1}^n q_j = 1,\ q_j \geq 0 \text{ for all } j = 1, \ldots, n \right\}$$

- 参与人 1 的策略 $\boldsymbol{p}$ 如果满足某一行 i 对应的 $p_i = 1$，则称该策略为**纯策略**（pure strategy），也可以写作 $\boldsymbol{e}^i$。同样，参与人 2 的纯策略满足某一列 j 对应的 $q_j = 1$，也可以写作 $\boldsymbol{e}^j$。

收益
Payoff

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

- 在二人零和博弈中，矩阵 A 的 (i,j) 要素 a_{ij} 是当参与人 1 选择纯策略 $\boldsymbol{e}^i$，同时参与人 2 选择纯策略 $\boldsymbol{e}^j$ 时，参与人 2 支付给参与人 1 的价值。也就是说，这时参与人 1 的收益是 a_{ij}，而参与人 2 的收益是 $-a_{ij}$。
- 如果参与人 1 选择策略 $\boldsymbol{p}$，参与人 2 选择策略 $\boldsymbol{q}$，则参与人 1 的**期望收益**是

$$\boldsymbol{p}A\boldsymbol{q} = \sum_{i=1}^{m}\sum_{j=1}^{n} p_i a_{ij} q_j$$

这里的正确写法是 $\boldsymbol{p}^\top A\boldsymbol{q}$，但是由于大多数情况下不会引起误解，所以简写为 $\boldsymbol{p}A\boldsymbol{q}$。

参与人 2 的期望收益是 $-\boldsymbol{p}A\boldsymbol{q}$。

Maximin/minimax 策略

- 策略 $\boldsymbol{p}$ 如果满足下列条件，则称它为参与人 1 在矩阵博弈 A 中的 maximin 策略

$$\min\{\underline{\boldsymbol{p}A\boldsymbol{q}} \mid \boldsymbol{q} \in \Delta^n\} \geq \underline{\min\{\boldsymbol{p}'A\boldsymbol{q} \mid \boldsymbol{q} \in \Delta^n\}} \quad \text{for all } \boldsymbol{p}' \in \Delta^m$$

① 对于参与人 1 的每一个策略 $\boldsymbol{p}'$，参与人 2 都会通过选择合适的 $\boldsymbol{q}$ 使参与人 1 的期望收益最小（也就是自身的期望支付最小）

② 而参与人 1 则通过选择合适的 $\boldsymbol{p}$ 在最小期望收益中实现最大化

- 策略 $\boldsymbol{q}$ 如果满足下列条件，则称它为参与人 2 在矩阵博弈 A 中的 minimax 策略

$$\max\{\boldsymbol{p}A\boldsymbol{q} \mid \boldsymbol{p} \in \Delta^m\} \leq \max\{\boldsymbol{p}A\boldsymbol{q}' \mid \boldsymbol{p} \in \Delta^m\} \quad \text{for all } \boldsymbol{q}' \in \Delta^n$$

对于参与人 2 来说，他是将自己的最大期望支付最小化

Maximin/minimax 策略

- 如果策略是有限的，那么肯定**存在** maximin 和 minimax 策略。
- 如何检验某个策略 $\boldsymbol{p}$ 或 $\boldsymbol{q}$ 是否使是 maximin/minimax 策略呢?
  - 以 $\boldsymbol{p}$ 为例，我们只需针对 $\boldsymbol{e}^j, j = 1, 2, \ldots, n$ 检验定义中的不等式即可
  - 这是因为任意的 $\boldsymbol{p}A\boldsymbol{q}$ 都是 $\boldsymbol{p}A\boldsymbol{e}^j$ 的线性组合，最小值只会出现在某个 $\boldsymbol{p}A\boldsymbol{e}^j$ 处：

$$\boldsymbol{p}A\boldsymbol{q} = \sum_{j=1}^{n} q_j\, \boldsymbol{p}A\boldsymbol{e}^j \quad \Rightarrow \quad \min\{\boldsymbol{p}A\boldsymbol{q} \mid \boldsymbol{q} \in \Delta^n\} = \min\{\boldsymbol{p}A\boldsymbol{e}^j \mid j = 1, 2, \ldots, n\}$$

    我们只需在这些 $\boldsymbol{p}A\boldsymbol{e}^j$ 中进行最大化。

  - 同样的逻辑也适用于 $\boldsymbol{q}$。因此，在寻找 maximin/minimax 策略时，参与人只需考虑对方的纯策略。

Minimax 定理

对任意矩阵博弈 A，存在一个实数 $v = v(A)$ 并满足下面的条件

a. 当且仅当策略 $\boldsymbol{p}$ 是参与人 1 的 maximin 策略时，它可以保证参与人 1 至少获得收益 v。

b. 当且仅当策略 $\boldsymbol{q}$ 是参与人 2 的 minimax 策略时，它可以保证参与人 2 最多只需支付 v。

- 我们称这个 $v = v(A)$ 为博弈 A 的**值**（value）。
- Maximin 和 minimax 策略分别是参与人 1 和参与人 2 的最优策略。

矩阵博弈的解

- “解”一个矩阵博弈意味着找出两个参与人的最优策略以及博弈的值。
- 以俾斯麦海海战博弈为例

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

盟军的北线策略对应的最小收益是 2，南线策略对应的最小收益是 1，因此 maximin 策略是**北线**（纯策略）。

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

日军的北线策略对应的最大支付是 2，南线策略对应的最大支付是 3，因此 minimax 策略是**北线**（纯策略）。

（北线，北线）策略组合使盟军的收益为 2，日军的支付为 2。因此博弈的值为 2。

鞍点

- 矩阵博弈 A 中的 (i,j) 要素 a_{ij} 如果满足下面的条件，则称其为**鞍点（saddle point）**：

$$a_{ij} \geq a_{kj} \quad \text{for all } k = 1,\ldots,m,$$
$$a_{ij} \leq a_{ik} \quad \text{for all } k = 1,\ldots,n,$$

即 a_{ij} 在列 j 中是最大值，在行 i 中是最小值。

- 如果 (i,j) 要素是鞍点，则参与人 1 可以获得的收益至少是 a_{ij}（因为在第 i 行中它是最小值），同时参与人 2 需要支付的最大值也是 a_{ij}（因为在第 j 列中它是最大值）。

因此 $a_{ij} = v(A)$，$\boldsymbol{e}^i$ 是参与人 1 的 maximin 策略，$\boldsymbol{e}^j$ 是参与人 2 的 minimax 策略。

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

用绘图法解 $2\times n$ 博弈

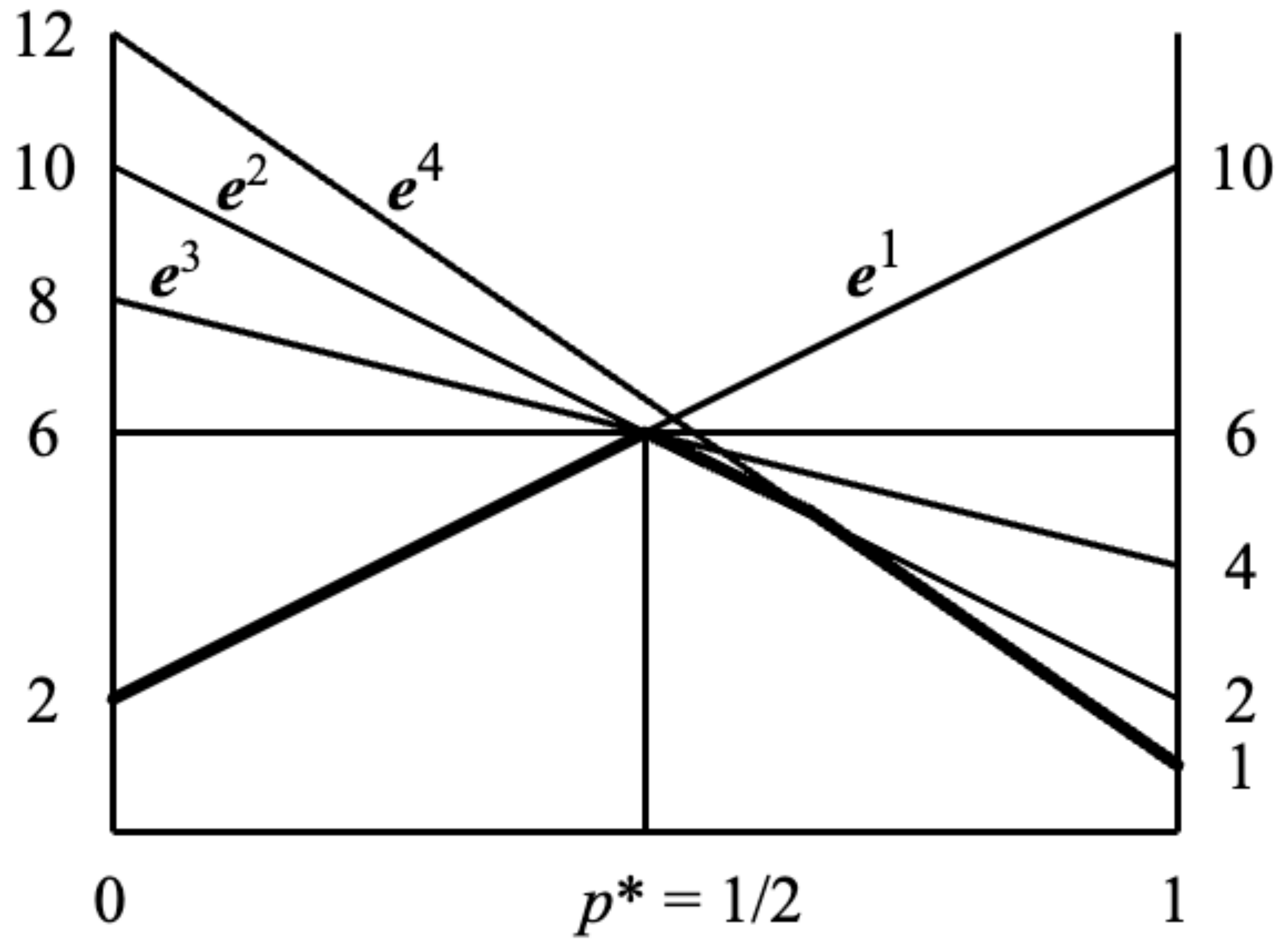
一个 $2 \times n$ 博弈

- 考虑下面的矩阵博弈

$$\begin{array}{cc} & \begin{array}{cccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 \end{array} \\ A = & \begin{pmatrix} 10 & 2 & 4 & 1 \\ 2 & 10 & 8 & 12 \end{pmatrix} \end{array}$$

- 令 $\boldsymbol{p} = (p, 1-p)$ 为参与人 1 的任意策略。
- 针对参与人 2 的每个纯策略，参与人 1 的期望收益是

$$\boldsymbol{p}A\boldsymbol{e}^1 = 10p + 2(1-p) = 8p + 2$$
$$\boldsymbol{p}A\boldsymbol{e}^2 = 2p + 10(1-p) = 10 - 8p$$
$$\boldsymbol{p}A\boldsymbol{e}^3 = 4p + 8(1-p) = 8 - 4p$$
$$\boldsymbol{p}A\boldsymbol{e}^4 = p + 12(1-p) = 12 - 11p$$



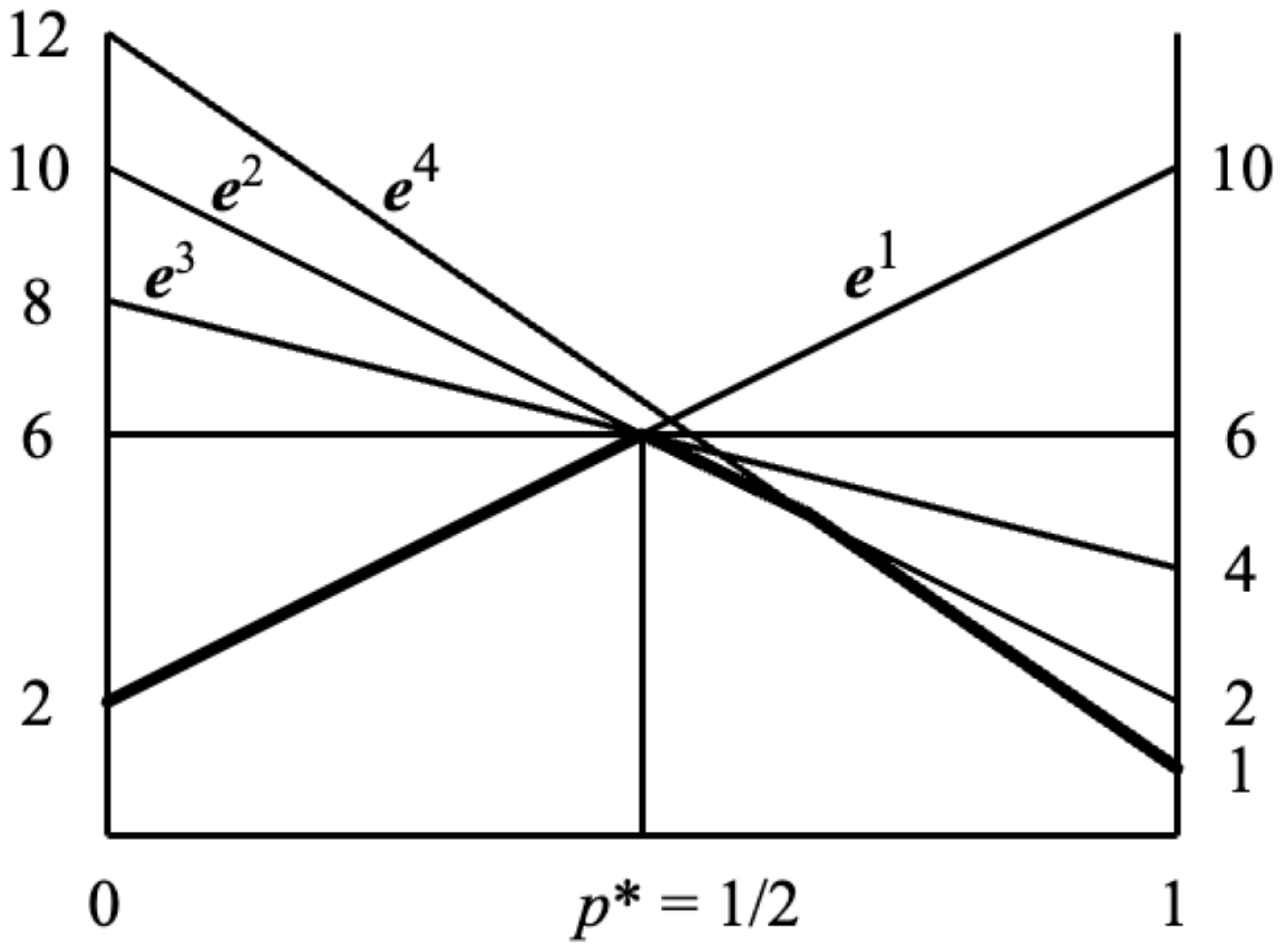
一个 $2 \times n$ 博弈

$$A = \begin{array}{c} \\ \\ \end{array}\begin{array}{cccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 \\ \left(\begin{matrix} 10 \\ 2 \end{matrix}\right. & \begin{matrix} 2 \\ 10 \end{matrix} & \begin{matrix} 4 \\ 8 \end{matrix} & \left.\begin{matrix} 1 \\ 12 \end{matrix}\right) \end{array}$$

- 图中的粗折线代表 $\min\{\boldsymbol{p}A\boldsymbol{q} \mid \boldsymbol{q} \in \Delta^n\}$ 的集合
- 参与人 1 的最优策略（maximin 策略）是

$$\boldsymbol{p}* = \left(\frac{1}{2}, \frac{1}{2}\right)$$

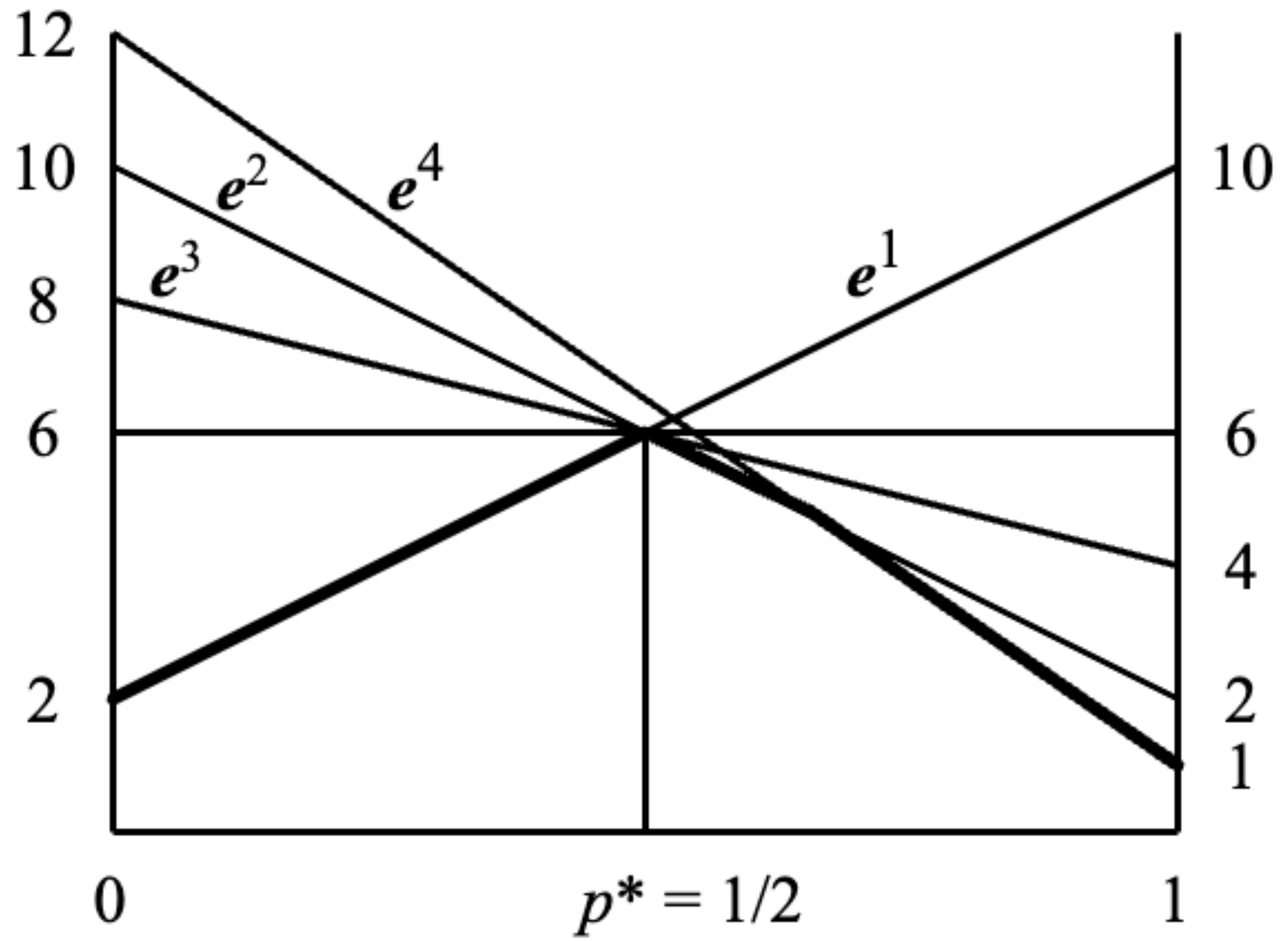
  这保证他至少能够获得收益 6。
- 根据 minimax 定理，可知博弈值为 $v(A) = 6$。



一个 $2 \times n$ 博弈

$$\begin{array}{cccccc} & & \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 \end{array}$$
$$A = \begin{pmatrix} 10 & 2 & 4 & 1 \\ 2 & 10 & 8 & 12 \end{pmatrix}$$

- 同样根据 minimax 定理，参与人 2 的 minimax 策略能够保证他的最大支付额为博弈值，即 6。
- 考虑参与人 2 的策略 $\boldsymbol{q} = (q_1, q_2, q_3, q_4)$。
- 右图中 $\boldsymbol{e}^4$ 对应的直线不通过点 $\left(\frac{1}{2}, 6\right)$，因此 $q_4 = 0$。即 minimax 策略满足 $\boldsymbol{q} = (q_1, q_2, q_3, 0)$。
- $\boldsymbol{q} = (q_1, q_2, q_3, 0)$ 是 $\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3$ 的线性组合，因此它对应的期望收益在图中应该是一条直线。此直线通过点 $\left(\frac{1}{2}, 6\right)$，且最大值为 6，因此可知其斜率为 0。



一个 $2 \times n$ 博弈

$$\begin{array}{cc} & \begin{array}{cccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 \end{array} \\ A = & \begin{pmatrix} 10 & 2 & 4 & 1 \\ 2 & 10 & 8 & 12 \end{pmatrix} \end{array}$$

- 策略 $\boldsymbol{q} = (q_1, q_2, q_3, 0)$ 在右图中对应的直线通过点 $(0,6)$ 和 $(1,6)$，因此可得以下两个方程

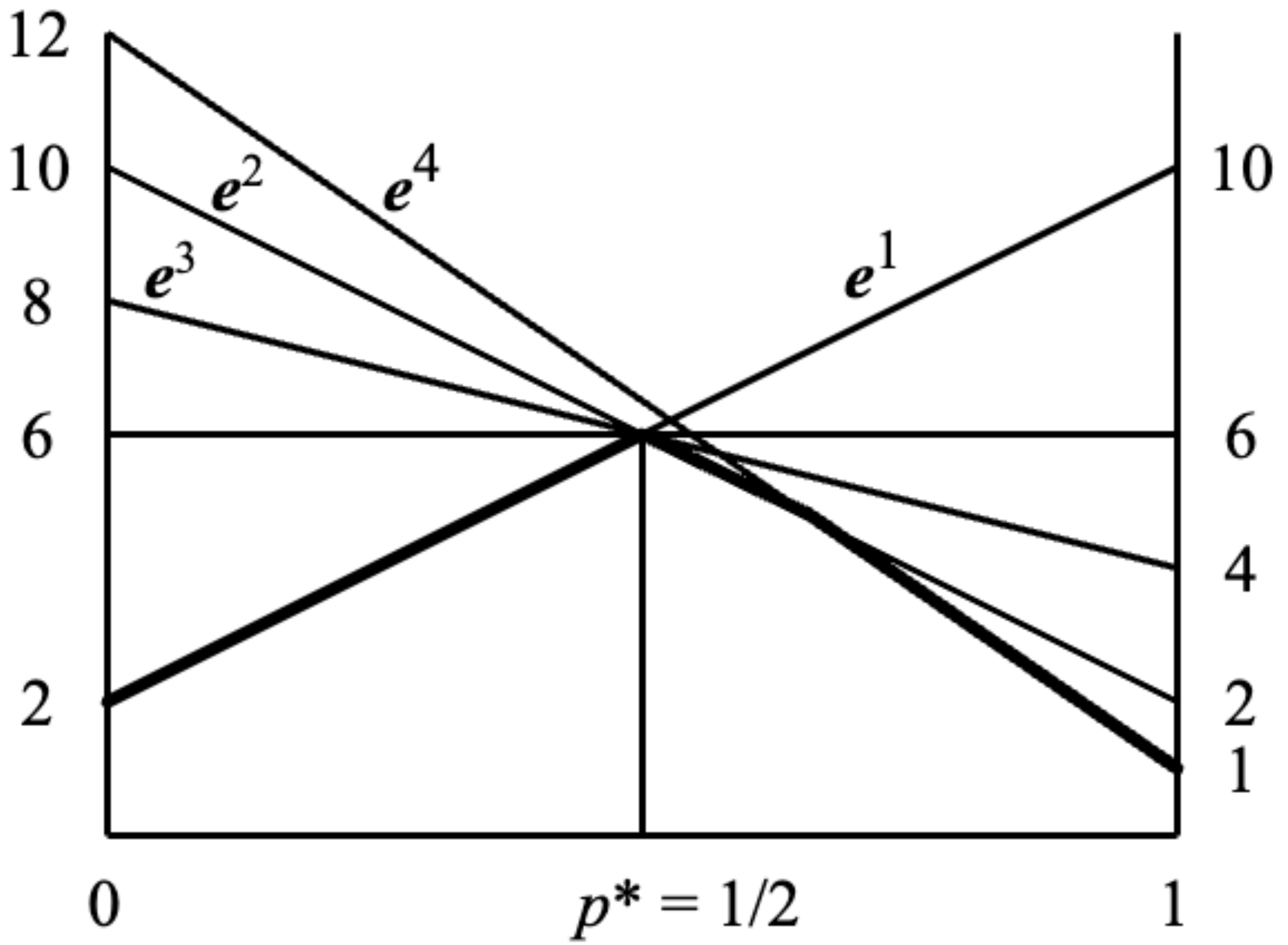
$$2q_1 + 10q_2 + 8q_3 = 6$$
$$10q_1 + 2q_2 + 4q_3 = 6$$

  消除 q_3 可得 $3q_1 - q_2 = 1$，因此参与人 2 的最优策略集合为

$$\left\{\boldsymbol{q} = (q_1, q_2, q_3, q_4) \in \Delta^4 \mid q_2 = 3q_1 - 1,\ q_4 = 0\right\}$$

- 根据 $0 \leq q_1, q_2, q_3, q_4 \leq 1$ 和 $q_1 + q_2 + q_3 = 1$，最优（minimax）策略集合可以进一步改写为

$$\left\{\boldsymbol{q} = (q_1, q_2, q_3, q_4) \in \mathbb{R}^4 \;\middle|\; \frac{1}{3} \leq q_1 \leq \frac{1}{2},\ q_2 = 3q_1 - 1,\ q_3 = 1 - q_1 - q_2,\ q_4 = 0\right\}$$



用绘图法解 $m \times 2$ 博弈

一个 $m \times 2$ 博弈

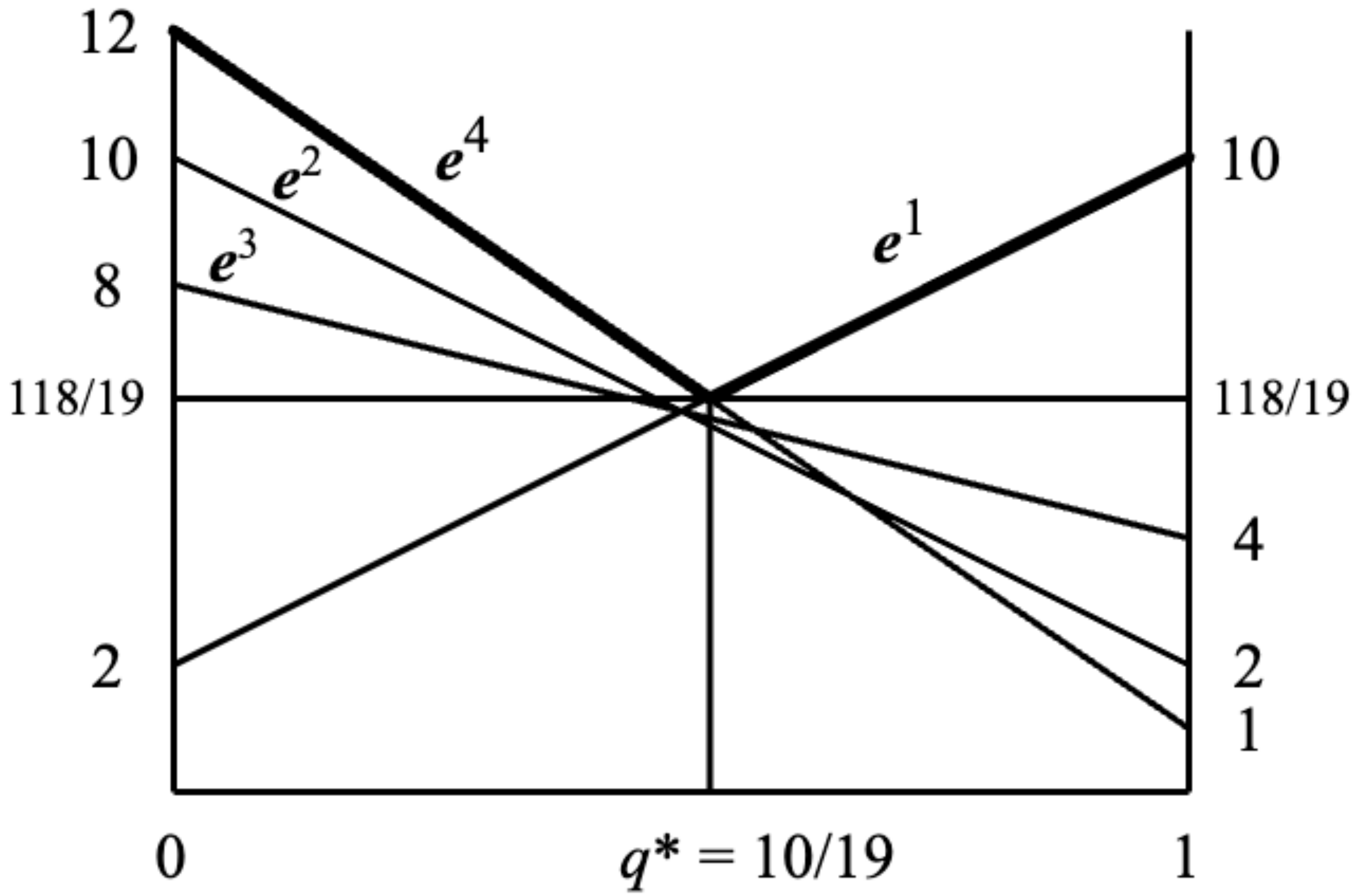
- 考虑下面的矩阵博弈

$$A = \begin{array}{c} e^1 \\ e^2 \\ e^3 \\ e^4 \end{array} \begin{pmatrix} 10 & 2 \\ 2 & 10 \\ 4 & 8 \\ 1 & 12 \end{pmatrix}$$

这是上一个例子中矩阵的转置

- 令 $\boldsymbol{q} = (q, 1-q)$ 为参与人 2 的任意策略。
- 针对参与人 1 的每个纯策略，参与人 2 的期望支付额是

$$\boldsymbol{e}^1 A \boldsymbol{q} = 10q + 2(1-q) = 8q + 2$$
$$\boldsymbol{e}^2 A \boldsymbol{q} = 2q + 10(1-q) = 10 - 8q$$
$$\boldsymbol{e}^3 A \boldsymbol{q} = 4q + 8(1-q) = 8 - 4q$$
$$\boldsymbol{e}^4 A \boldsymbol{q} = q + 12(1-q) = 12 - 11q$$



一个 $m \times 2$ 博弈

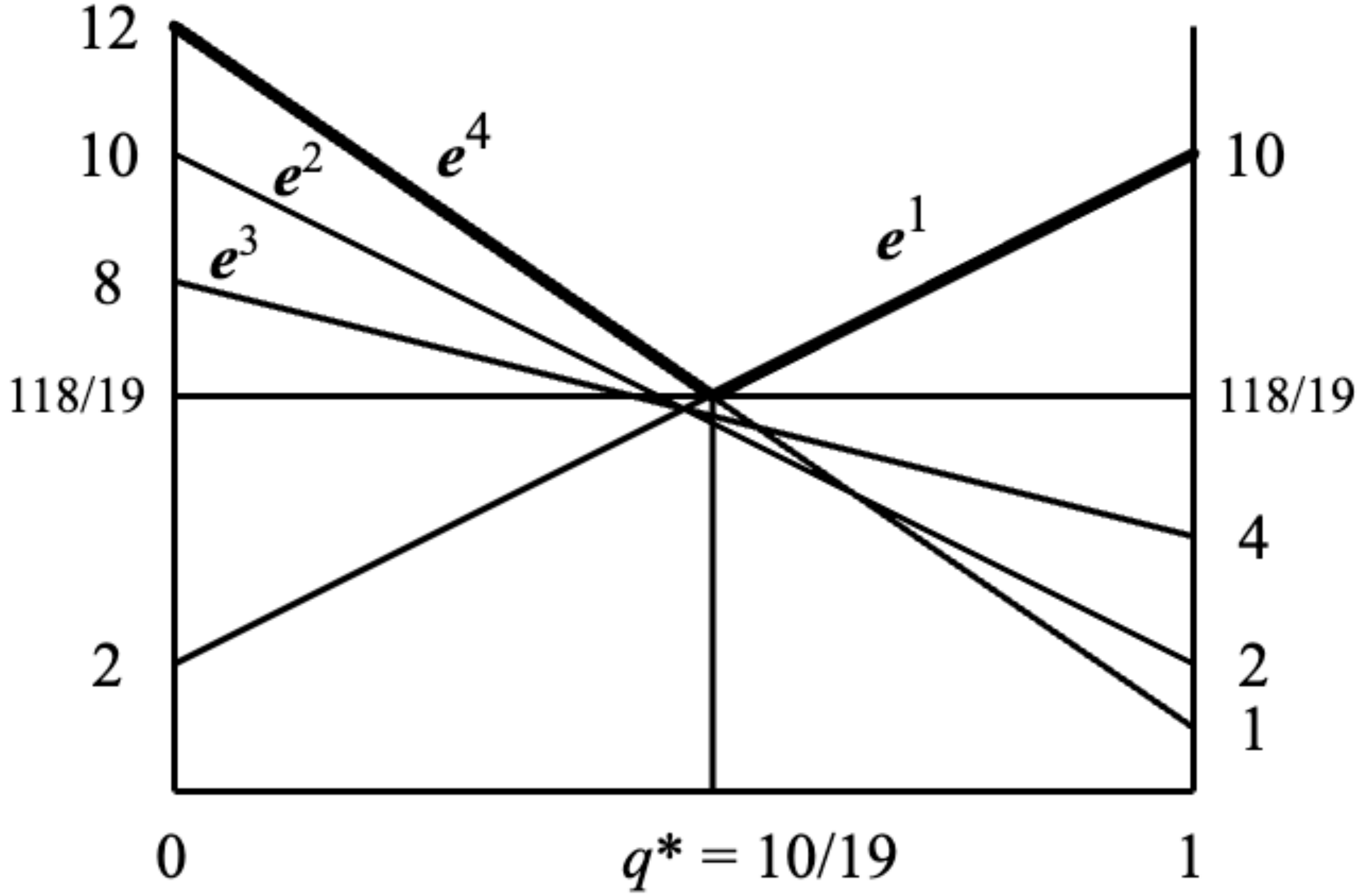
$$A = \begin{array}{c} e^1 \\ e^2 \\ e^3 \\ e^4 \end{array} \begin{pmatrix} 10 & 2 \\ 2 & 10 \\ 4 & 8 \\ 1 & 12 \end{pmatrix}$$

- 参与人 2 的最大支付集合对应右图中的粗折线，其最小值为 118/19，发生在 $q^* = 10/19$ 处。因此参与人 2 的 minimax 策略为

$$\boldsymbol{q}^* = \left(\frac{10}{19}, \frac{9}{19}\right)$$

  博弈值为 118/19。

- 参与人 1 的最右策略 $\boldsymbol{p} = (p_1, p_2, p_3, p_4)$ 是 $\boldsymbol{e}^1$ 和 $\boldsymbol{e}^4$ 的线性组合，因此 $p_2 = p_3 = 0$。由 minimax 定理可知，参与人 1 的最优策略能够保证其收益至少为 118/19，因此最优策略对应的直线的斜率为 0。



一个 $m \times 2$ 博弈

$$A = \begin{array}{c} e^1 \\ e^2 \\ e^3 \\ e^4 \end{array} \begin{pmatrix} 10 & 2 \\ 2 & 10 \\ 4 & 8 \\ 1 & 12 \end{pmatrix}$$

- 参与人 1 的最优策略对应的直线通过点 $\left(0,\ \frac{118}{19}\right)$ 和 $\left(1,\ \frac{118}{19}\right)$，因此可得方程

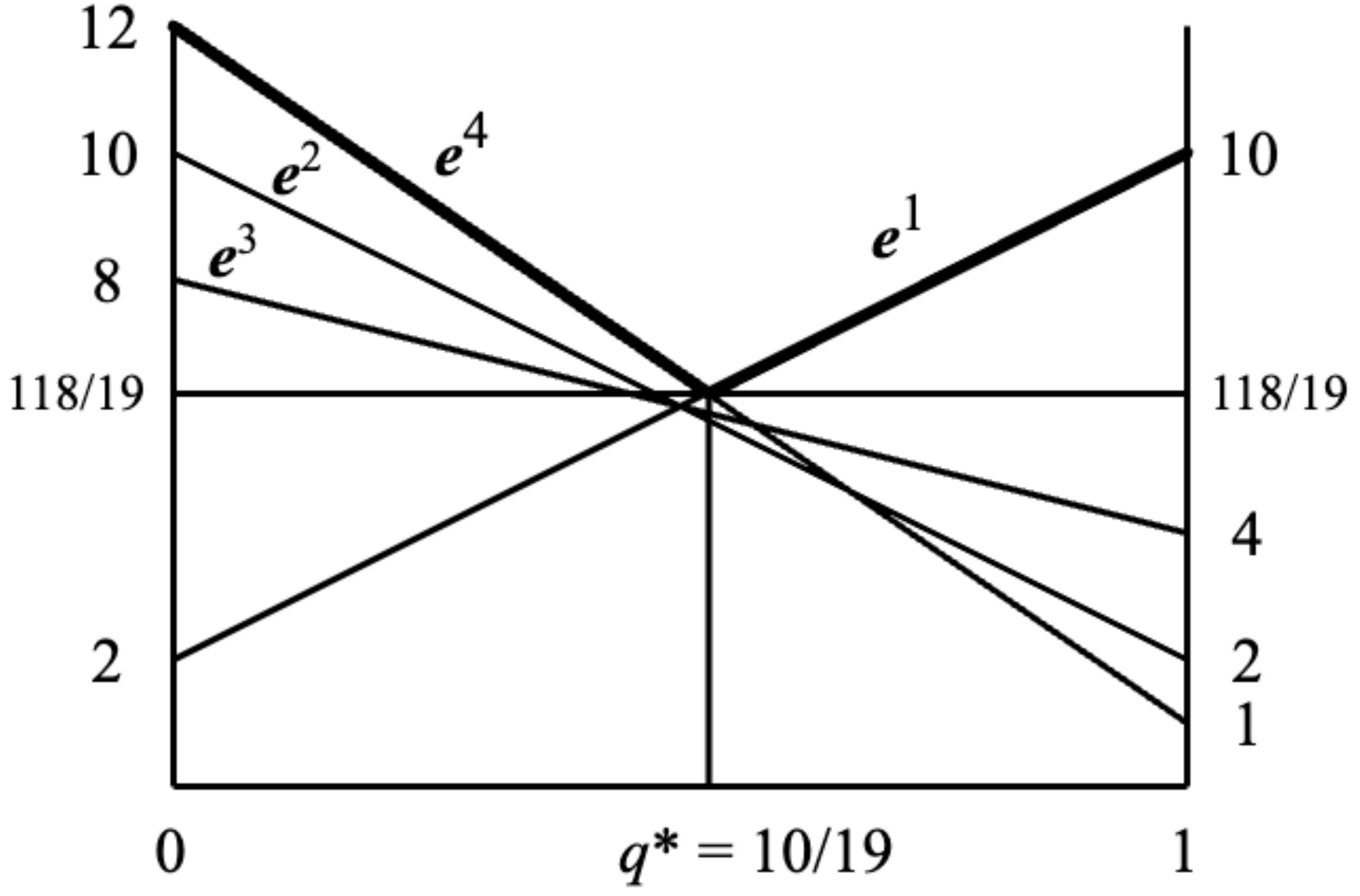
$$2p_1 + 12p_4 = \frac{118}{19}$$

$$10p_1 + p_4 = \frac{118}{19}$$

  其解为 $(p_1, p_4) = \left(\frac{11}{19}, \frac{8}{19}\right)$。

- 因此参与人 1 的最优（maximin）策略是唯一的，即

$$\boldsymbol{p}^* = \left(\frac{11}{19},\ 0,\ 0,\ \frac{8}{19}\right)$$



严格优势（strict domination）

严格优势

- 对于一个参与人的两个策略 $\boldsymbol{p}^1$ 和 $\boldsymbol{p}^2$，如果无论对方怎样选择，$\boldsymbol{p}^1$ 带来的收益永远高于 $\boldsymbol{p}^2$ 带来的收益，则称 $\boldsymbol{p}^1$ **严格优于（strictly dominates）**$\boldsymbol{p}^2$，或 $\boldsymbol{p}^2$ **严格劣于（strictly dominated by）**$\boldsymbol{p}^1$。如果一个策略严格劣于另一策略，则称前者为严格劣势策略。

- 考虑矩阵博弈

$$
\begin{array}{cc}
 & \begin{array}{ccccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 & \boldsymbol{e}^5 \end{array} \\
A = & \begin{pmatrix} 10 & 2 & 5 & 1 & 6 \\ 2 & 10 & 8 & 12 & 9 \end{pmatrix}
\end{array}
$$

  这里参与人 2 的策略 $\boldsymbol{e}^5$ 就严格劣于 $\boldsymbol{e}^3$。 注意参与人 2 的收益是矩阵中数字的负数

- 在最优策略中，严格劣势策略的概率为零。因此我们可以事先剔除严格劣势策略。

混合策略与纯策略间的优劣

$$\begin{array}{cc} & \begin{array}{ccccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 & \boldsymbol{e}^5 \end{array} \\ A = & \begin{pmatrix} 10 & 2 & 5 & 1 & 6 \\ 2 & 10 & 8 & 12 & 9 \end{pmatrix} \end{array}$$

- 我们继续考虑参与人 2 的策略 $\boldsymbol{e}^3$。
- 在第一行中，$\boldsymbol{e}^3$ 的收益 5 介于 $\boldsymbol{e}^1$ 的收益 10 和 $\boldsymbol{e}^2$ 的收益 2 之间；而在第二行中，$\boldsymbol{e}^3$ 的收益 8 也介于 $\boldsymbol{e}^1$ 的收益 2 和 $\boldsymbol{e}^2$ 的收益 10 之间。那么是否存在一个 $\boldsymbol{e}^1$ 和 $\boldsymbol{e}^2$ 的混合策略能够严格优于 $\boldsymbol{e}^3$ 呢?
- 如果选择 $\boldsymbol{e}^1$ 的概率为 α，选择 $\boldsymbol{e}^2$ 概率为 $1-\alpha$，则混合策略 $(\alpha, 1-\alpha, 0, 0, 0)$严格优于纯策略 $\boldsymbol{e}^3$ 的条件是

$$\alpha\begin{pmatrix}10\\2\end{pmatrix}+(1-\alpha)\begin{pmatrix}2\\10\end{pmatrix}=\begin{pmatrix}8\alpha+2\\10-8\alpha\end{pmatrix}<\begin{pmatrix}5\\8\end{pmatrix}$$

向量间的大小关系：
$\boldsymbol{a} < \boldsymbol{b}$ 代表所有要素
均满足 $a_i < b_i$

解不等式可得 $\frac{1}{4} < \alpha < \frac{3}{8}$。在此范围内，参与人 2 在最优策略中选择 $\boldsymbol{e}^3$ 的概率也为零。

剔除严格劣势策略

$$A = \begin{array}{c} \\ \\ \end{array}\begin{array}{c} \begin{array}{ccccc} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 & \boldsymbol{e}^4 & \boldsymbol{e}^5 \end{array} \\ \begin{pmatrix} 10 & 2 & 5 & 1 & 6 \\ 2 & 10 & 8 & 12 & 9 \end{pmatrix} \end{array}$$

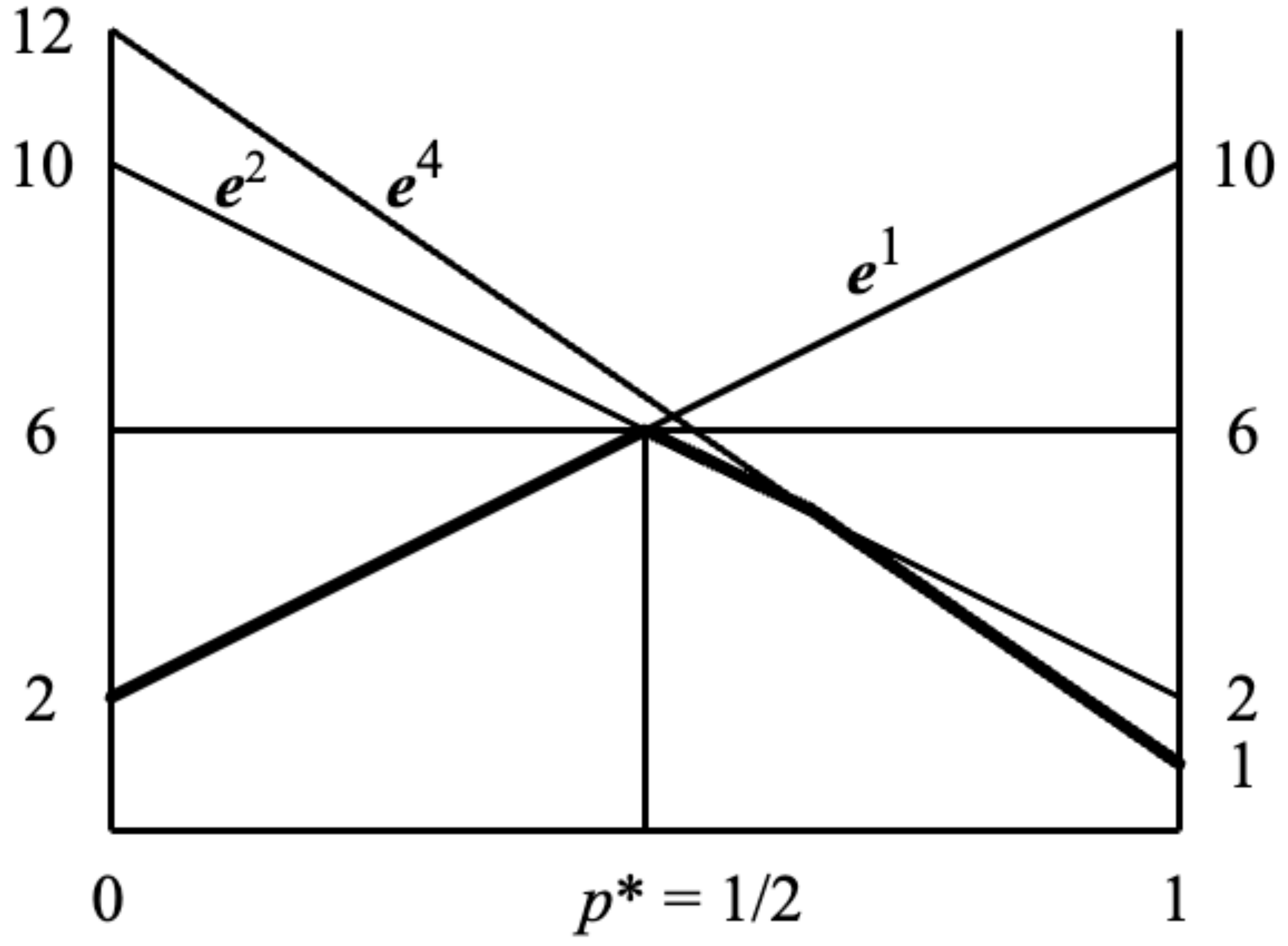
- 将严格劣势策略 $\boldsymbol{e}^5$ 和 $\boldsymbol{e}^3$ 剔除后再进行图形分析，可以让分析过程简化。右图即为剔除这两个策略后的情形。
- 右图中可知，博弈值为 6，参与人 1 的最优策略为

$$\boldsymbol{p}^* = \left(\frac{1}{2}, \frac{1}{2}\right)$$

- 参与人 2 的最优策略显然不包含 $\boldsymbol{e}^4$，且依然为通过 (0, 6) 和 (1, 6) 的直线，解方程组

$$2q_1 + 10q_2 = 6$$
$$10q_1 + 2q_2 = 6$$

可得 $\boldsymbol{q}^* = \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0\right)$。



Example 2.1

- 考虑下面的 3×3 矩阵博弈

$$A = \begin{pmatrix} 6 & 0 & 2 \\ 0 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

- 参与人 1 的策略 $\boldsymbol{e}^3$ 严格劣于策略 $\boldsymbol{p} = \left(\frac{7}{12}, \frac{5}{12}, 0\right)$。剔除矩阵 A 的第三行可得

$$B = \begin{pmatrix} 6 & 0 & 2 \\ 0 & 5 & 4 \end{pmatrix}$$

- 此时，参与人 2 的策略 $\boldsymbol{e}^3$ 严格劣于策略 $\boldsymbol{q} = \left(\frac{1}{4}, \frac{3}{4}, 0\right)$。剔除矩阵 B 的第三列可得

$$C = \begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}$$

- 怎么解矩阵博弈 C 呢?
  - 是否存在鞍点?
  - 能不能用图形分析?