有限ゼロ和二人ゲーム

有限ゼロ和二人ゲームは、最も簡単な標準形ゲームです。次の3つのルールから成り立ちます。(1)プレイヤーの数は二人 (2)ゲームの結果について、二人の利得の和は常に0 (3)各プレイヤーのとれる戦略は有限個

例
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & & & \vdots \\ \vdots & & a_{ij} & \vdots \\ a_{n1} & \cdots & \cdots & a_{nm} \end{pmatrix} \qquad A = \begin{pmatrix} 4 & -1 & -2 \\ -3 & 0 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

n×m 利得行列

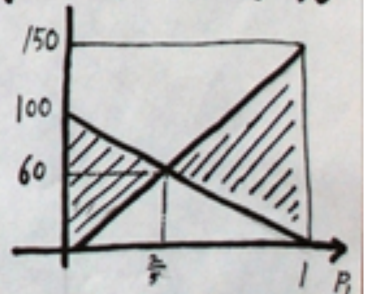
ミニマックス原理

プレイヤー1,2(P_1,P_2)が有限ゼロ和ゲームを行います。P_1が戦略i, P_2がjを選んだ際のP_1の利益を例のような利得行列で表せます。このとき、P_1は利得を最大にする最大化プレイヤーで、P_2は利得を最小にする最小化プレイヤーと考えることができます。今二人の利益は完全に対立しており単純な最大化最小化原理ではない新しい行動原理が必要となります。そこで登場するのがこのミニマックス原理です。例として上のAのような利得行列を考えます。P_1が$(i=1)$の戦略を用いた場合、P_1が一番好ましくないのはP_2が$\min(4,-1,-2)=-2$となる$(j=3)$の戦略を用いた場合です。同様にP_1が$(i=2)$を用いた場合は$\min(-3,0,3)=3$で$(j=1)$が、$(i=3)$の時は$\min(3,1,2)=1$で$(j=2)$が好ましくありません。この内、P_1に一番利益となる$\max(-2,-3,1)=1$となる$(i=3)$をP_1は選ぶべきだ、この考え方がミニマックス原理です。P_2の立場からこの考えを使うと、P_2が選ぶべき戦略がP_1と一致します。この時、1はゲームの均衡点です。

混合戦略

$$B = \begin{array}{c|cc|c} & & & P_1 \\ & 0 & 150 & 1 \\ & 100 & 0 & 2 \\ \hline P_2 & 1 & 2 & \end{array}$$

先程のようにミニマックス原理では二人の戦略が一致しない時もあります。このような場合必要になるのが混合戦略です。Bの行列をP_1の利得行列とすると先程のような均衡点は存在しません。よって混合戦略で均衡点を見つけることになります。P_1が戦略1,2を用いる確率をp_1, p_2、P_2が戦略1,2をq_1, q_2とすると、P_1が獲得する利益の期待値は $E_1(p, q) = 100 \times p_2 \times q_1 + 150 \times p_1 \times q_2$ となります。ここで $p_1+p_2=1$, $q_1+q_2=1$, $0 \leqq p_1, q_1 \leqq 1$ であることから、図のようなP_1の利得グラフがかけます。



最小化プレイヤーであるP_2はグラフの一番低くなる値になるようなqをえらびます。その内でP_1は最大の値のとれるp_1として$p_1=\frac{2}{5}$, $p_2=\frac{3}{5}$を選択するのが良いと分かります。以上よりP_1は戦略1,2をそれぞれ$\frac{2}{5}$, $\frac{3}{5}$の確率で選択すべきと分かり、同様の議論でP_2も戦略1,2をそれぞれ$\frac{3}{5}$, $\frac{2}{5}$の確率で選択すべきであり、さらにお互い利益60で均衡すると分かります。

定理（良い戦略）

P_1, P_2と立場が全く同じならば、P_1が得る利得の期待値が一番大きくなる戦略pを立てた場合、P_2にとっても戦略pは自分の利益を最大にする戦略となり、加えて最大にすることができる戦略は必ずpだけになります。

有限ゼロ和二人ゲームの知識を使ってポーカーについて考える。具体的にはどのような戦略をとれば利得の期待値が最大になるかを考える。まず、実際のポーカーのルールは多彩かつ複雑ゆえ、以下のようにルールを単純化する。

(i) プレイヤーは二人 (ii) 手札はお互い最初に配られた5枚のみ、交換はしない。

(iii) 用いる戦略は a, b $(a>b)$ なる数が与えられたと仮定して以下の3通りとする。

(1) a (高い額)をビッド(かける) (2) 始め b (低い額)をビッドし、相手が a をビッドしたらコール(a をビッド)

(3) 始め b をビッドして、相手が a をビッドしたら降りる。

(iv) プレイヤーは同時にビッドする。(v) 同額ビッドの場合、カードが強い方の勝ち、ビッド分の利得を得る。負けた方はビッド分利得を失う。降りた場合はそのプレイヤーが負け、利得を相手に払う。(vi) カードの強さは $[0,1]$ 上の一様分布とし、数が大きい方が強い。

(vi)について…本来カードの強さは ${}_{52}C_5$ 通りで離散的であるが、一番強いカードを1、弱いカードを0、他は順番に並べると、見た目には密になり、連続としても精度の高い解答が得られると考え、そのように仮定した。プレイヤー1がカード s_1 を持ち、プレイヤー2がカード s_2 を持っているとき、P_1 の戦略を i_{s_1} $(i_{s_1}=1,2,3)$, P_2 の戦

略を j_{s_2} ($j_{s_2}=1,2,3$) とすると、P_1 が得る利得は

$$\mathrm{sgn}(s_1-s_2)=\begin{cases} + & (s_1>s_2) \\ - & (s_1<s_2) \end{cases},\ \varphi_+(i,j),\ \varphi_-(i,j)$$

i \ j	1	2	3
1	a	a	b
2	a	b	b
3	-b	b	b

$\varphi_+(i,j)$

i \ j	1	2	3
1	-a	-a	b
2	-a	-b	-b
3	-b	-b	-b

$\varphi_-(i,j)$

厳密に

連続か

次に、P_1 の混合戦略を $\rho=\begin{pmatrix}\rho_1(s)\\ \rho_2(s)\\ \rho_3(s)\end{pmatrix}$

P_2 の混合戦略を $\sigma=\begin{pmatrix}\sigma_1(s)\\ \sigma_2(s)\\ \sigma_3(s)\end{pmatrix}$

とし、P_2 が得る利得の期待値を

$K(\rho,\sigma)=\sum_{j=1}^{3}\int_0^1 r_j(\rho,s)\sigma_j(s)ds$ と表される。

$$r_1(\rho,s)=\int_0^s(-a\rho_1(t)-a\rho_2(t)-b\rho_3(t))dt$$
$$+\int_s^1(a\rho_1(t)+a\rho_2(t)-b\rho_3(t))dt \cdots ①$$

$$r_3(\rho,s)=\int_0^s(b\rho_1(t)-b\rho_2(t)-b\rho_3(t))dt$$
$$+\int_s^1(b\rho_1(t)+b\rho_2(t)+b\rho_3(t))dt \cdots ③$$

$\varphi_{\mathrm{sgn}(s_1-s_2)}(i_{s_1}, j_{s_2})$ と表せる。ここで、

は次の行列で表される。

は $s_1=s_2$ の場合が必要であるが s_1, s_2 は

つ、a, b は有限なので、そこは無視する。

(sはカードの強さ、$s\in[0,1]$, $\rho_i(s)$ はカードs を持っているときに純戦略(i)を選ぶ確率
$\rho_1+\rho_2+\rho_3=1$ ($\forall s$))

…上と同様

$K(\rho,\sigma)$ とすると、

ここで $r_j(\rho,s)$ は

$$r_2(\rho,s)=\int_0^s(-a\rho_1(t)-b\rho_2(t)-b\rho_3(t))dt$$
$$+\int_s^1(a\rho_1(t)+b\rho_2(t)-b\rho_3(t))dt \cdots ②$$

であり、P_1のカードの強さがs、混合戦略ρを用いて、P_2が純戦略jを用いたときのP_2が得る期待値を表している。

この$K(\rho,\sigma)$を最小にするρを求めたい。

($K(\rho,\sigma)$が$\sigma(s)$に関して最小になるための条件は$r_j(\rho,s)$がjについて最小値をとらなければ$\sigma_j(s)=0$ということである。従って$\rho=\sigma$の時、最小化されるという良い戦略の判断基準は$r_j(\rho,s)$がjについて最小値をとらなければ$\rho_j(s)=0$となることである。しかもそのような時に限られる。⊛)

まず、あるsに対して$\rho_2(s)>0$と仮定する。このとき$\min_j r_j(\rho,s)=r_2(\rho,s)$

①,②より$(a-b)\left(\int_0^s \rho_2(t)dt-\int_s^1 \rho_2(t)dt\right)+2b\int_s^1 \rho_3(t)dt \leqq 0$ —④ となる。次に$\rho_2(s)>0$を満たすsの上限をs_0とする。すると、積分の連続性から$s=s_0$に対しても④が成り立つ。$s>s_0$に対しては$\rho_2(s)=0$なので、④において$\int_s^1 \rho_2(t)dt=0$である。よって④は、

$(a-b)\int_0^{s_0}\rho_2(t)dt+2b\int_{s_0}^1 \rho_3(t)dt \leqq 0$と表せる。

$\rho_2(s_0)>0$ かつ $\rho_2(t)\geqq 0$ より $\int_0^{s_0}\rho_2(t)dt>0$ よって第1項は正である。しかし明らかに第2項も非負より矛盾。従って任意の $s\in[0,1]$ に対して $\rho_2(s)=0$、ゆえに $\rho_3(s)=1-\rho_1(s)$ $(0\leqq\rho_1(s)\leqq 1)$ 次に $\rho_1(u)\neq 0,1$ となる点 u について考える。この u を中間点といい、$\rho_1(u)\neq 0$ のとき $\min_j r_j(\rho,u)=r_1(\rho,u)$ となり、$\rho_1(u)\neq 1$ $(\rho_3(u)\neq 0)$ のとき $\min_j r_j(\rho,u)=r_3(\rho,u)$ となる。よって、このような点 u のどんな近傍においても $r_1(\rho,u')\leqq r_3(\rho,u')$ となる点 u' と、$r_1(\rho,u'')\geqq r_3(\rho,u'')$ となる点 u'' が存在するので r_j の連続性より点 u について $r_3(\rho,u)-r_1(\rho,u)=0$ が成り立つ。①,③より $(a+b)\int_0^u\rho_1(t)dt-(a-b)\int_u^1\rho_1(t)dt+2b\int_u^1(1-\rho_1(t))dt=0$

$\Leftrightarrow (a+b)\left(\int_0^u\rho_1(t)dt-\int_u^1\rho_1(t)dt\right)+2b(1-u)=0$ —④ さらに2つの中間点 u', u'' $(u'<u'')$ について考え、上の式の u に u', u'' を代入してその差を求めると、$2(a+b)\int_{u'}^{u''}\rho_1(t)dt-2b(u''-u')=0$

従って $\frac{1}{u''-u'}\int_{u'}^{u''}\rho_1(t)dt=\frac{a}{a+b}$ —⑤ 即ち、2つの中間点 u', u'' の間での $\rho_1(s)$ の平均値は、$\frac{a}{a+b}$ となるので、$u'\leqq t\leqq u''$ の区間で、恒等的に $\rho_1(t)=1$ となることも $\rho_1(t)=0$ となることもない。もしそうなれば平均は1又は0になるからである。ゆえにこの区間 $[u', u'']$ の間には、中間点が致るところ密に存在する。よって⑤の積分の連続性によって、$u'\leqq t\leqq u''$ となる全ての t において、$\rho_1(t)=\frac{b}{a+b}$ となる。ここで、中間点が必ず存在することを示しておく。常に $\rho_1(t)=0$ とすると、$r_1(\rho,0)=-b$

$r_3(\rho,0)=b$ となる。よって $r_1(0)<r_2(0)$ と
常に $\rho_3(t)=0$ とする場合も同様。従って
最大、最小がある。$\bar{u}'$ を中間点の最小値、
において $\rho_1(t)=\frac{b}{a+b}$，ところで $\bar{u}''<t\leq 1$
$\rho_1(t)=1$ または $\rho_1(t)=0$，前の結論より
$\rho_1(t)=1$，$0\leq t\leq \bar{u}'$ においても恒等的に
すると r_3-r_1 即ち ☆ の左辺にこれを

$$r_3-r_1=2(a+b)\cdot 0-2b=-2b<0$$

ばならず、矛盾。$\rho_1(t)=1$ とすると r_3-r_1
$r_3>r_1$ より常に $\rho_1(t)=0$ とならなければ
従って、$\bar{u}'=0$，最後に $\bar{u}''$ の値を求

$$-(a+b)\int_0^1 \rho_1(t)\,dt+2b=0,$$

$$\int_0^1 \rho_1(t)\,dt=\bar{u}''\cdot\frac{b}{a+b}+(1-\bar{u}'')\cdot 1=$$

なり、$\rho_3(0)=0$ となるが、$\rho_3(0)=1$ に矛盾。
中間点 u は必ず1つは存在するので
$\bar{u}''$ を中間点の最大値とすると、$\bar{u}'\leq t\leq \bar{u}''$
には中間点が存在しないので、恒等的に
$\rho_1(t)=0$ はあり得ない。よって $\bar{u}''\leq t\leq 1$ で
$\rho_1(t)=0$ または $\rho_1(t)=1$ となる。$\rho_1(t)=0$ と
代入して、
r_2-r_1 より常に $\rho_1(t)=1$ とならなけれ

$$=2(a+b)\cdot 1-2b=2a>0$$ であるが、

ならず、これも矛盾。
めたい。☆ の u に 0 を代入して

$$\int_0^1 \rho_1(t)\,dt=\frac{2b}{a+b}$$

$1-\frac{a}{a+b}\bar{u}''$ より $1-\frac{a}{a+b}\bar{u}''=\frac{2b}{a+b}$　$\bar{u}''=\frac{a-b}{a}$

以上により、求める混合戦略は

$$\rho_1(s)=\begin{cases}\dfrac{a}{a+b} & \left(0\leqq s\leqq \dfrac{a-b}{a}\right)\\ 1 & \left(\dfrac{a-b}{a}<s<1\right)\end{cases}$$

$$\rho_2(s)=0\quad(0\leqq s\leqq 1)$$

$$\rho_3(s)=\begin{cases}\dfrac{b}{a+b} & \left(0\leqq s\leqq \dfrac{a-b}{a}\right)\\ 0 & \left(\dfrac{a-b}{a}<s\leqq 1\right)\end{cases}$$

となることがわかる。

ゲーム理論

岩田雄大
片山友貴
坂井俊亮