

Einführung in die Spieltheorie

Joachim Hermisson

June 9, 2016

Das Gefangenendilemma (Prisoner's dilemma)

Zwei Personen werden verdächtigt, eine Bank ausgeraubt zu haben, aber abgesehen von illegalem Waffenbesitz kann die Staatsanwaltschaft ihnen bisher nichts nachweisen. Damit sie sich nicht absprechen können, werden sie in der Untersuchungshaft strikt getrennt und einzeln verhört. Die Staatsanwaltschaft bietet beiden eine Kronzeugenregelung an. Der Straferlass für den Kronzeugen richtet sich nach der Bedeutung seiner Aussage im Verfahren. Vollen Straferlass bekommt er genau dann, wenn seine Aussage für die Anklagevertreter entscheidend und alternativlos ist. Wenn die Anklage sich alternativ auf mehrere Aussagen (oder andere Beweise) stützen kann, darf ein geständiger Angeklagter immerhin noch auf eine Reduktion seiner Strafe hoffen.

Das erwartete Strafmaß für illegalen Waffenbesitz bewegt sich bei einem Jahr. Ein nicht geständiger Bankräuber muss mit 10 Jahren Haft rechnen, ein geständiger Räuber mit sieben Jahren. Was sollte ein Verdächtiger in einer solchen Situation tun?

Im der Terminologie der Spieltheorie haben beide Akteure je zwei (reine) Strategien zur Verfügung:

- untereinander zu kooperieren (*cooperate*), also: nicht aussagen;
- die Kooperation zu verweigern (*defect*), also: gestehen und den anderen bezichtigen.

Das Resultat jeder dieser Strategien hängt aber vom Verhalten des anderen ab. Man kann dies in einer sogenannten Auszahlungsmatrix (*payoff matrix*) zusammenfassen:

$$\begin{array}{cc} & C & D \\ \begin{array}{c} C \\ D \end{array} & \left(\begin{array}{cc} -1 & -10 \\ 0 & -7 \end{array} \right), & (1) \end{array}$$

wobei C und D für *cooperate* und *defect* stehen. Die eigene Strategie entscheidet über die Zeile des Matrixeintrags, die des Mitspielers (Mitangeklagten) über die Spalte. Über die Wahl des Mitspielers habe ich keine Information. Dennoch erscheint eine rationale Entscheidung beim Gefangenendilemma einfach: unabhängig von der Strategie des anderen werde ich immer mit der Entscheidung für *defect* (also verraten/bezichtigen) besser

fahren. Da dies für beide gilt, sehen sie sich alsbald für 7 Jahre in Haft. Und das ist das Dilemma: die rationale Entscheidung beider Akteure führt zu einem suboptimalen Ergebnis. Wenn man die Summe der Jahre in Haft als Gesamtergebnis sieht, dann ist dieses sogar das schlechteste mögliche.

Bemerkungen Das Gefangenendilemma hat eine große Bedeutung für die Wirtschaftsethik: Es zeigt, dass die Maximierung des persönlichen Nutzens durch alle Akteure nicht unbedingt den Gesamtnutzen für die Gesellschaft maximiert (es kann diesen sogar minimieren). Das abstrakte Spiel beschreibt dabei eine Situation, die sich erstaunlich häufig und in ganz verschiedenen Zusammenhängen des "wirklichen Lebens" ergeben kann. Beispiele sind:

- Das Gefangenendilemma gilt als starkes Argument gegen Kronzeugenregelungen, da die Kalkulation unabhängig von der tatsächlichen Schuld der Verächtigen ist.
- In vielen praktischen Beispielen geht es um Ressourcenverbrauch zur Schädigung / Bedrohung von Mitbewerbern:
 - Werbung konkurrierender Unternehmen bei gesättigtem Markt (Folge USA 1971: Tabak-Industrie befürwortet Werbeverbot für Zigaretten).
 - Patentschlachten in der IT-Branche mit zu diesem Zweck erworbenen Patentpaketen.
 - Rüstungswettlauf und Gleichgewicht des Schreckens.
 - Dominanz- und Revierkämpfe in der Tierwelt (und nicht nur dort).
- Steuerhinterziehung und Schwarzarbeit – wo immer das Risiko einer Bestrafung gering ist und massenhafte Hinterziehung den Staat ruiniert.
- Weite Bereiche der Umweltpolitik, z.B. Regeln zum CO₂ Ausstoß zwischen Ländern oder Fangverbote bei Überfischung.
- Kartellrecht zur Verhinderung von Absprachen bei Oligopolen (z.B. OPEC Ölfördermengen): mangelnde Kooperation nützt Einzelnen, aber wenn niemand sich an die Absprachen hält, werden die Preise ruiniert.
- Versteigerungen mit Bieterwettbewerben (ebay, staatliche Versteigerungen, z.B. BAWAG in Österreich).

In der Wirtschaftsethik begründen Gefangenendilemmata oft staatliches Handeln: Durch geeignete Rahmenbedingungen (z.B. Umweltgesetze) sollen unerwünschte Dilemmata vermieden werden: Stärkung von Kooperation. Andererseits sollen erwünschte Dilemmata herbeigeführt werden, z.B. im Kartellrecht: Stärkung von Wettbewerb.

Grundlegende Definitionen

Im Gefangenendilemma haben beide Spieler genau zwei Verhaltensweisen zur Verfügung: “Aussagen” oder “Nicht Aussagen”. Allgemein können es auch mehr Verhaltensweisen sein. Man nennt diese Verhaltensweisen in der Spieltheorie auch **reine Strategien**. Die Auszahlungsmatrix \mathbf{A} eines symmetrischen 2-Personen Spiels mit N reinen Strategien (*normal form game*) ist wie folgt definiert:

$$\mathbf{A} = (a_{ij})_{i,j=1,\dots,N} \quad (2)$$

wobei a_{ij} die Auszahlung für die reine Strategie i ist, wenn der Gegenspieler die reine Strategie j spielt. Bei einem symmetrischen Spiel haben beide Spieler die gleiche Auszahlungsmatrix. Im allgemeineren asymmetrischen Fall können sie sich unterscheiden (Bi-Matrix Spiele). Beide Spieler können dabei auch völlig unterschiedliche Strategien zur Verfügung haben.

- Neben den reinen Strategien, die immer die gleiche Wahl treffen, gibt es sogenannte **gemischte Strategien**, die sich mit Wahrscheinlichkeit p_i für die reine Strategie i entscheiden. Eine gemischte Strategie kann durch einen Wahrscheinlichkeitsvektor $\mathbf{p} = (p_1, p_2, \dots, p_N)^t$ mit $\sum_i p_i = 1$ dargestellt werden. Allgemein ergibt sich der *payoff* einer gemischten Strategie \mathbf{p} gegen eine Strategie \mathbf{q} als Skalarprodukt

$$\mathbf{p} \cdot \mathbf{Aq} \quad (3)$$

- Eine Strategie heißt **dominant**, wenn sie unter allen Strategien immer den höchsten *payoff* hat, unabhängig von der Strategie des Mitspielers. Man sieht leicht, dass dafür hinreichend und notwendig ist, dass die Strategie über alle anderen reinen Strategien dominiert. Die Strategie i dominiert deshalb, wenn für alle j und für alle $k \neq i$ immer $a_{ij} > a_{kj}$ ist. Beim Gefangenendilemma dominiert D über jede andere Strategie.
- Eine Strategie heißt symmetrisches **Nash Gleichgewicht**, wenn kein Spieler durch den Wechsel der Strategie seinen *payoff* vergrößern kann. Für $\mathbf{q} \neq \mathbf{p}$ ist \mathbf{p} ein Nash Gleichgewicht, wenn

$$\mathbf{p} \cdot \mathbf{Ap} \geq \mathbf{q} \cdot \mathbf{Ap} \quad (4)$$

Ein symmetrisches Nash Gleichgewicht ist also immer die “beste Antwort auf sich selbst”. Wenn sogar $\mathbf{p} \cdot \mathbf{Ap} > \mathbf{q} \cdot \mathbf{Ap}$ gilt, heißt das Nash Gleichgewicht *strikt* (einzige beste Antwort auf sich selbst). Im Gefangenendilemma ist D (gestehen bzw. verraten) ein striktes Nash Gleichgewicht. Allgemein kann man zeigen, dass es bei einem 2-Personen Spiel mit N reinen Strategien immer mindestens ein (nicht unbedingt striktes) Nash Gleichgewicht gibt (siehe Hofbauer/Sigmund 1998). Für den asymmetrischen Fall erweitert man die Definition wie folgt: ein Paar von Strategien ist ein Nash-Gleichgewicht, wenn kein Spieler durch die einseitige Änderung seiner Strategie seinen *payoff* verbessern kann.

- Es ist offensichtlich, dass eine dominante Strategie immer ein striktes Nash Gleichgewicht ist – und dass die Implikation andersherum nicht gilt.

Fragen

Die Auszahlungsmatrix \mathbf{A} eines 2-Personen Spiels mit 2 reinen Strategien A und B ist

$$\mathbf{A} = \begin{array}{c} \begin{array}{cc} C & D \\ C & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ D & \end{array} \end{array} \quad (5)$$

1. a) Was sind die Bedingungen dafür, dass C bzw. D ein (strikt) Nash Gleichgewicht ist? b) Wie viele Nash Gleichgewichte mit reinen Strategien kann es in einem Spiel mit N reinen Strategien geben?
2. Wie würde die Auszahlungsmatrix des Spiels "Schere – Stein – Papier" aussehen? Was wäre hier ein Nash Gleichgewicht? Ist es strikt?
3. Hausaufgabe 1: Kann ein gemischter (nicht reiner) Zustand dominant sein? Kann ein gemischter Zustand ein striktes Nash Gleichgewicht sein? [Nein und nein. Für ein gemischtes Nash Gleichgewicht ist für alle $p_i > 0$ immer $(\mathbf{A}\mathbf{p})_i = c$. Damit sind alle Linearkombinationen der entsprechenden reinen Zustände gleich gute Antworten auf \mathbf{p} .]
4. Hausaufgabe 2: Zeige, dass es in einem Spiel mit zwei reinen Strategien immer mindestens ein Nash Gleichgewicht gibt. Berechne das Gleichgewicht. Wann ist es strikt?
5. Spiele "Schere – Stein – Papier" gegen den Computer. Wer hat zuerst 100 Siege? <http://www.nytimes.com/interactive/science/rock-paper-scissors.html>

Zu Hausaufgabe 2:

$$\begin{aligned} \mathbf{q} \cdot \mathbf{A}\mathbf{p} &= pqa + (1-p)qb + p(1-q)c + (1-p)(1-q)d \\ &= pq(a+d-b-c) + p(c-d) + q(b-d) + d \end{aligned}$$

p ist ein Nash Gleichgewicht, wenn (bei p festgehalten) die Strategie $q = p$ die Auszahlung maximiert. Da der *payoff* linear in q ist, liegt das Maximum entweder am Rand (bei $q = 0$ oder bei $q = 1$) oder die *payoff*-Funktion hat Steigung null. Für $d \geq b$ ist $p(= q) = 0$ ein Nash Gleichgewicht (strikt für $d > b$). Für $a \geq c$ ist $p = 1$ ein Nash Gleichgewicht (strikt für $a > c$). Für $d < b$ und $a < c$ ist $p = (b-d)/(b-d-a+c)$ ein (nicht striktes) Nash Gleichgewicht, das gegen jede Strategie q den gleichen *payoff* hat.

Evolutionäre Spiele

Die klassische Spieltheorie geht davon aus, dass Spieler "rational" sind und auch mit "rationalen" Gegen- (oder Mit-) Spielern rechnen können. Dies ist bereits für menschliches Verhalten fraglich. Für Anwendungen in der Biologie ist es unbrauchbar. Strategien sollten sich hier durch Evolution innerhalb einer Population selbst herausbilden. In einer evolutionären Dynamik setzt man den mittleren *payoff* eines Individuums

in einer Population mit der erwarteten Anzahl seiner Nachkommen gleich (biologische *Fitness*). Welche Strategie ein Individuum spielt ist genetisch vorgegeben und wird an die Nachkommen vererbt. Deshalb wird die Frequenz einer Strategie in der Population genau dann zunehmen, wenn ihr mittlerer *payoff* höher ist als der mittlere *payoff* der gesamten Population. Durch Mutationen können neue Strategien in einer Population auftreten. Die Theorie evolutionärer Spiele geht auf John Maynard Smith zurück und wurde u.a. von Karl Sigmund und Josef Hofbauer in Wien zu einer anspruchsvollen mathematischen Theorie entwickelt.

Tauben und Falken

Warum kämpfen Tiere innerhalb einer Art meist eher zurückhaltend und nach Regeln, die schwere Verletzungen vermeiden? – Es ist keine Frage, dass dies zur Erhalt der Art beiträgt, aber das ist evolutionär kein gutes Argument. Innerhalb einer Population mit zurückhaltenden Individuen könnte aggressives Verhalten für den Einzelnen dennoch erfolgreich sein und sich so durchsetzen. Eine spieltheoretische Erklärung stammt von John Maynard Smith, dem Begründer der evolutionären Spieltheorie.

Beim *Hawk and Dove* Spiel kann ein Individuum entweder eine eskalierende Falken-Strategie F oder eine de-eskalierende Tauben-Strategie T einnehmen kann. Die *payoff* Matrix ist wie folgt:

$$\begin{array}{c} T \quad F \\ T \left(\begin{array}{cc} b/2 & 0 \\ b & (b-c)/2 \end{array} \right) \\ F \end{array} \quad (6)$$

Zwei Tauben teilen sich die Beute, ein Falke wird sie gegen eine Taube alleine behalten, zwei Falken kämpfen und zahlen dafür Kosten. Wichtig ist vor allem der Fall $c > b$, bei dem "Falke – Falke" (oder "*defect – defect*") für alle Beteiligten das schlechteste Resultat ist. Ein anderer Name für ein Spiel dieses Typs ist *chicken* (nach einem Spiel, bei dem junge Helden aus den Vorstädten mit Autos aufeinander zurasen – wer ausweicht ist ein *chicken*).

Strategien können rein ("Falke" (F) oder "Tauben" (T)) oder gemischt sein. Für eine einfache evolutionäre Dynamik nehmen wir zunächst an, dass es nur zwei Strategien in der Population gibt: eine vorherrschende *Wildtyp* Strategie \mathbf{p} und eine Mutanten-Strategie \mathbf{q} . Wenn der *payoff* für eine seltene Mutanten-Strategie höher ist als der Wildtyp *payoff*, dann wird sie sich in der Population etablieren (man sagt auch: sie kann *invadieren*), anderenfalls stirbt sie wieder aus.

Fragen

1. Betrachte zunächst das Taube-Falke Spiel als gewöhnliches Spiel zwischen zwei Spielern. Wie lautet das Nash Gleichgewicht? ($x = b/c$ Wahrscheinlichkeit, als Falke zu reagieren)

2. Betrachte nun das evolutionäre Spiel. Nimm an, dass der Wildtyp die reine F Strategie ist. Welche anderen Strategien können dann invadieren? Welche Strategien können umgekehrt von F oder T invadiert werden?
3. Wie wird durch solch ein Spiel (teilweise) Deeskalation erklärt?

Evolutionäre Stabilität

Betrachten wir allgemein eine Population, die anfänglich nur aus Individuen mit Wildtyp-Strategie \mathbf{p} besteht. Eine andere Mutanten-Strategie \mathbf{q} tritt mit sehr geringen Frequenz ϵ in der Population auf. Diese andere Strategie wird anwachsen, wenn ihre (mittlere) Fitness (= *payoff*) in der Population höher als die Fitness von \mathbf{p} ist:

$$\mathbf{q} \cdot \mathbf{A}((1 - \epsilon)\mathbf{p} + \epsilon\mathbf{q}) > \mathbf{p} \cdot \mathbf{A}((1 - \epsilon)\mathbf{p} + \epsilon\mathbf{q}) \quad (7)$$

für alle hinreichend kleinen ϵ . Wir können auch fragen, ob es überhaupt eine Strategie \mathbf{q} gibt, die invadieren kann. Dies führt zum Konzept der evolutionären Stabilität.

Definition

1. Eine Strategie \mathbf{p} , die gegen die Invasion jeder anderen Strategie stabil ist, heißt eine *evolutionär stabile Strategie*, kurz *ESS*. Wir leiten dafür zwei Bedingungen ab
 - a) Gleichgewichtsbedingung

$$\mathbf{q} \cdot \mathbf{A}\mathbf{p} \leq \mathbf{p} \cdot \mathbf{A}\mathbf{p} \quad \forall \mathbf{q} \quad (8)$$

- b) Stabilitätsbedingung

$$\text{wenn } \mathbf{q} \neq \mathbf{p} \text{ und } \mathbf{q} \cdot \mathbf{A}\mathbf{p} = \mathbf{p} \cdot \mathbf{A}\mathbf{p} \Rightarrow \mathbf{q} \cdot \mathbf{A}\mathbf{q} < \mathbf{p} \cdot \mathbf{A}\mathbf{q} \quad (9)$$

2. Wenn in der Stabilitätsbedingung nur gilt $\mathbf{q} \cdot \mathbf{A}\mathbf{q} = \mathbf{p} \cdot \mathbf{A}\mathbf{q}$, nennt man dies auch ein *schwaches ESS*.

Die Gleichgewichtsbedingung ist genau die eines Nash Gleichgewichts. Die Stabilitätsbedingung kommt hinzu für den Fall, dass das Nash Gleichgewicht nicht strikt ist: Sie besagt, dass \mathbf{p} besser gegen \mathbf{q} abschneidet als \mathbf{q} selbst. Mit der Definition von Dominanz von oben haben wir also die Relation

$$\text{dominant} \Rightarrow \text{strikt Nash} \Rightarrow \text{ESS} \Rightarrow \text{schwaches ESS} \Rightarrow \text{Nash}. \quad (10)$$

Für das Spiel der Tauben und Falken bekommen wir mit $\mathbf{p} = (1 - b/c, b/c)$ und $\mathbf{q} = ((1 - q), q)$

$$\begin{aligned} \mathbf{p} \cdot \mathbf{A}\mathbf{q} &= \frac{(1 - q)(c - b)b}{2c} + \frac{(1 - q)b^2}{c} + \frac{q(b - c)b}{2c} = \frac{b^2 + bc - 2bqc}{2c} \\ \mathbf{q} \cdot \mathbf{A}\mathbf{q} &= \frac{(1 - q)^2 b}{2} + q(1 - q)b + \frac{q^2(b - c)}{2} = \frac{bc - q^2 c^2}{2c} \end{aligned}$$

und daher

$$\mathbf{p} \cdot \mathbf{A}\mathbf{q} - \mathbf{q} \cdot \mathbf{A}\mathbf{q} = \frac{(b - qc)^2}{2c} > 0 \quad \text{für } \mathbf{q} \neq \mathbf{p}. \quad (11)$$

Das Nash Gleichgewicht ist daher evolutionär stabil.

Fragen

1. Gibt es immer ein ESS? [nein, z.B. die *payoff* Matrix mit allen Einträgen gleich]
2. Gibt es immer zumindest ein schwaches ESS? [nein, z.B. Schere – Stein – Papier Variante mit Gewinn 1 und Verlust > 1]

Replikatorgleichung (replicator equation)

Bisher haben wir nur implizite Annahmen über die evolutionäre Dynamik gemacht: die Frequenz einer Strategie in der Population nimmt zu/ab wenn ihr *payoff* höher/niedriger ist als der mittlere *payoff* der anderen Strategien in der Population. Dies ist ausreichend, um das Anwachsen (oder nicht) einer seltenen Mutanten-Strategie in einer Wildtyp Population zu beurteilen. Für Populationen, in denen mehrere Strategien gleichzeitig auftreten können, müssen wir die Dynamik genauer bestimmen. Dies ist dann besonders einfach, wenn wir uns auf Typen mit reinen Strategien beschränken. Wie oben stellt man sich vor, dass der *payoff* einer Strategie einer biologischen Fitness w_i entspricht, also die erwartete Anzahl überlebender Nachkommen für die nächste Generation bestimmt. Dann breiten sich erfolgreiche Strategien durch natürliche Selektion in einer Population aus. Wenn x_i die Frequenz ist, mit der eine (reine) Strategie i in einer Population gespielt wird, dann gilt

$$\dot{x}_i = x_i(w_i - \bar{w}) \quad (12)$$

$$= x_i \left(\sum_{j=1}^n x_j a_{ij} - \sum_{i,j=1}^n a_{ij} x_i x_j \right) \quad (13)$$

$$= x_i \left((\mathbf{Ax})_i - \mathbf{x} \cdot \mathbf{Ax} \right). \quad (14)$$

Dies ist die sogenannte Replikatorgleichung. $\mathbf{x} = (x_1, \dots, x_n)$ ist der Frequenzvektor und $\bar{w} = \mathbf{x} \cdot \mathbf{Ax}$ ist die mittlere Fitness in der Population. Alle Strategien mit einer größeren Fitness als \bar{w} wachsen an. Die Dynamik erfüllt $\sum_i \dot{x}_i = 0$ und erhält deshalb die Normierung des Frequenzvektors, $\sum_i x_i = 1$.

- Allgemeiner kann man auch die Replikatordynamik zwischen m beliebigen (gemischten oder reinen) Strategien in einer Population betrachten. Wenn $\mathbf{p}_1, \dots, \mathbf{p}_m$ die Vektoren für diese Strategien sind, dann definiert $\mathbf{p}_i \cdot \mathbf{Ap}_j$ den *payoff* von \mathbf{p}_i gegen \mathbf{p}_j und

$$\dot{x}_i = x_i \left(\sum_{j=1}^m x_j \mathbf{p}_i \cdot \mathbf{Ap}_j - \sum_{i,j=1}^m \mathbf{p}_i \cdot \mathbf{Ap}_j x_i x_j \right) \quad (15)$$

die Dynamik der Frequenz x_i der i -ten Strategie. Es ist einfach zu sehen, dass sich dies auf (13) reduziert, wenn wir die n reinen Strategien $\mathbf{p}_i = \mathbf{e}_i$ verwenden. Wir werden uns im Folgenden auf diesen Fall beschränken.

- Man sieht, dass die Dynamik unter der Replikatorgleichung unverändert bleibt, wenn man eine Konstante zu jedem Eintrag a_{ij} in einer Spalte der Auszahlungsmatrix addiert (also für alle i bei festgehaltenem j , $a_{ij} \rightarrow a_{ij} + c_j$). Man kann diese Invarianz dazu nutzen, alle Einträge a_{ii} auf der Hauptdiagonalen auf Null zu bringen - oder auch alle Einträge der letzten Zeile.

Strategien und Zustände

Der Frequenzvektor \mathbf{x} beschreibt einen Zustand der Population als Mischung reiner Strategien. Ein solcher Misch-Zustand einer Population ist in vieler Hinsicht analog zu einer gemischten Strategie eines (einzelnen) Gegenspielers: Man kann \mathbf{x} quasi als gemischte Strategie auf der Ebene der Population auffassen. Analog zu den (*normal form*) Spielen kann man auch in Populationen Nash Gleichgewichte und evolutionär stabile Zustände definieren:

- Ein Zustand \mathbf{x} ist ein Nash Gleichgewicht der Replikatorgleichung mit Fitness-Matrix \mathbf{A} wenn für alle Frequenzvektoren \mathbf{y} gilt

$$\mathbf{y} \cdot \mathbf{A}\mathbf{x} \leq \mathbf{x} \cdot \mathbf{A}\mathbf{x}.$$

- Ein Zustand \mathbf{x} heißt evolutionär stabil wenn

$$\mathbf{x} \cdot \mathbf{A}\mathbf{y} > \mathbf{y} \cdot \mathbf{A}\mathbf{y}$$

für alle Zustände \mathbf{y} in einer Umgebung von \mathbf{x} gilt. (Nachdem man Zustände in einer Umgebung von \mathbf{x} immer als $(1 - \epsilon)\mathbf{x} + \epsilon\mathbf{y}'$ darstellen kann, entspricht diese Bedingung genau der Gleichgewichts- und Stabilitätsbedingung für evolutionär stabile Strategien, wie oben definiert.)

- Daneben kann man noch stationäre Punkte (oder evolutionäre Gleichgewichte) \mathbf{x} definieren, bei denen

$$\dot{\mathbf{x}} = \mathbf{0}$$

gilt.

Man kann zeigen, dass ein evolutionär stabiler Zustand \mathbf{x} immer auch ein asymptotisch stabiler stationärer Punkt der Replikatordynamik ist (siehe zB Hofbauer & Sigmund 1998, *Evolutionary Games and Population Dynamics*). Das heißt: wenn wir in der Nähe des Punktes starten (in einer ϵ Umgebung), dann werden wir unter der Dynamik immer nach \mathbf{x} konvergieren. Wie bereits diskutiert sind evolutionär stabile Zustände immer auch Nash Gleichgewichte. Für das Verhältnis von Nash Gleichgewichten und (asymptotisch stabilen oder nicht stabilen) stationären Punkten geben die Antworten auf die folgenden Fragen Auskunft:

Fragen

1. Sind alle Nash Gleichgewichte stationäre Punkte? (Ja, denn $(\mathbf{Ax})_i = \mathbf{e}_i \cdot \mathbf{Ax} \leq \mathbf{x} \cdot \mathbf{Ax}$. Damit ist für alle x_i immer $\dot{x}_i \leq 0$. Da $\sum_i \dot{x}_i = 0$ (Normierung), ist $\dot{x}_i \equiv 0$ die einzige Lösung.)
2. Sind umgekehrt alle stationären Punkte Nash Gleichgewichte? (Nein, z.B. sind alle monomorphen Zustände stationär, aber nicht unbedingt Nash).

Replikatorndynamik für 2 reine Strategien

Für nur 2 reine Strategien C und D mit Auszahlungs/Fitness-Matrix \mathbf{A} wie in Gleichung (5) reduziert sich das System auf eine einzige DGL für $x = x_C$:

$$\dot{x} = x(1-x)[(a-b-c+d)x+b-d]. \quad (16)$$

Damit kann man leicht eine Klassifizierung der Dynamik erhalten, indem man das Verhalten an den beiden monomorphen Zuständen $x = 0$ und $x = 1$ betrachtet (jeweils mit Diagramm):

- C dominiert: $a > c$ und $b > d$. (Für ein einziges ESS am Rand reicht es, wenn eine dieser Ungleichungen strikt ist).
- Entsprechend D dominiert: $a < c$ und $b < d$.
- A und B beide stabil (und Nash): $a > c$ und $b < d$.
- C und D beide instabil: $a < c$ und $b > d$. Inneres Nash Gleichgewicht = ESS bei

$$x = \frac{d-b}{a-b-c+d}. \quad (17)$$

- C und D neutral: $a = c$ und $b = d$.

Fragen

1. Hausaufgabe: Betrachte das verallgemeinerte Schere - Stein - Papier Spiel mit Auszahlungsmatrix

$$\mathbf{A} = \begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix}. \quad (18)$$

Wie lautet das Nash Gleichgewicht und wann ist es ein ESS? (Antwort: Bei innerem Nash Gleichgewicht muss sein: $-a_2p_2 + b_3p_3 = b_1p_1 - a_3p_3 = -a_1p_1 + b_2p_2$ oder $(b_3 + a_3)p_3 = b_1p_1 + a_2p_2$ und damit $((a_3 + b_3)b_2 + a_2)p_2 = (a_1(a_3 + b_3) + b_1((a_3 + b_3) - 1))p_1$, etc. $\mathbf{p} = C(b_2b_3 + b_2a_3 + a_2a_3, b_3b_1 + b_3a_1 + a_1a_3, b_1b_2 + b_1a_2 + a_1a_2)$ und ESS bei $b_i > a_{i+1}$. Rechnung direkt über ESS Bedingung möglich.)

Spieltheorie und Lotka-Volterra Modelle

Die allgemeine Lotka-Volterra Gleichung beschreibt die Interaktion von n Arten, wobei die Wachstumsrate jeder Art eine lineare Funktion der Häufigkeit jeder Art ist:

$$\dot{y}_i = y_i \left(r_i + \sum_{j=1}^n b_{ij} y_j \right). \quad (19)$$

Hier ist y_i die Häufigkeit von Spezies i mit Wachstumsrate r_i , solange alle Artenzahlen klein sind. Die Konstanten b_{ij} geben die Auswirkung der Gegenwart von Individuen der Art j auf die Wachstumsrate der Art i an. Sie können positiv oder negativ sein. Typischerweise ist $b_{ii} < 0$ (Dichteregulation). Für $i \neq j$ bedeutet $b_{ij} < 0$, dass Art j einen negativen Einfluss auf das Wachstum von i hat (z.B. der Räuber für die Beute); $b_{ij} > 0$ zeigt einen positiven Einfluss ab (z.B. Beute für den Räuber).

Es zeigt sich, dass das Lotka-Volterra System mit n Arten äquivalent zu einer Replikatorgleichung mit $n + 1$ Typen ist:

$$\dot{x}_i = x_i \left(\sum_{j=1}^{n+1} a_{ij} x_j - \sum_{i,j=1}^{n+1} a_{ij} x_i x_j \right). \quad (20)$$

Um das zu sehen, zeigen wir erst die Quotientenregel:

$$\left(\frac{\dot{x}_i}{x_j} \right) = \frac{\dot{x}_i x_j - x_i \dot{x}_j}{x_j^2} = \frac{x_i}{x_j} \left(\sum_{k=1}^{n+1} a_{ik} x_k - \sum_{k=1}^{n+1} a_{jk} x_k \right). \quad (21)$$

Dann definieren wir

$$y_i = \frac{x_i}{x_{n+1}}.$$

Ohne Einschränkung können wir zudem annehmen, dass $a_{n+1,i} = 0$ ist (allgemein transformiere $a_{i,j} \rightarrow a_{i,j} - a_{n+1,j}$). Dann ist mit $y_{n+1} = 1$:

$$\dot{y}_i = y_i \sum_{k=1}^{n+1} a_{ik} y_k x_{n+1} = y_i \left(a_{i,n+1} + \sum_{k=1}^n a_{ik} y_k \right) x_{n+1}. \quad (22)$$

Als letzten Schritt entfernen wir den Faktor x_{n+1} . Da er alle y_i in gleicher Weise beeinflusst, ist dies durch eine Zeittransformation $t \rightarrow \tau = t \cdot x_{n+1}$ möglich. (Es ändert sich nur die Geschwindigkeit mit der die Trajektorien der Populationsdynamik durchlaufen werden, aber nicht die Trajektorien (und Gleichgewichte) selbst). Mit $r_i = a_{i,n+1}$ und $b_{ij} = a_{ij}$ erhalten wir nun das Lotka-Volterra System [Beweis Hofbauer 1981].

Ergänzung: Bi-Matrix Spiele

Bisher hatten wir immer angenommen, dass die Spielsituation für beide Spieler immer symmetrisch ist: sie haben die gleichen Strategien zur Verfügung und den gleichen *pay-off*. Dies ist keineswegs notwendig. Im allgemeinen Fall müssen wir zwischen Spielern in

Position 1 und 2 unterscheiden. Zu jeder Spiel-Position gehört dann ein eigener Satz von Strategien mit resultierenden *payoffs*, die sich in je einer Matrix (**A** und **B**) zusammenfassen lassen. Wenn z.B. **A** $n \times m$ dimensional ist, dann ist **B** $m \times n$ dimensional. Eine gemischte Strategie für den Spieler in Position 1 (Position 2) wird dann durch einen n dimensionalen Wahrscheinlichkeitsvektor **p** (bzw. einen m dimensionalen W.-Vektor q) definiert. Der *payoff* ergibt sich als $\mathbf{p} \cdot \mathbf{A}\mathbf{q}$, bzw. als $\mathbf{q} \cdot \mathbf{B}\mathbf{p}$. Ein Nash Gleichgewicht ist dann durch ein Paar von Strategien **p** und **q** gegeben, die wechselseitig die beste Antwort aufeinander sind. Ein Beispiel ist die folgende Situation:

- John und Oskar spielen Poker mit nur zwei Karten (z.B. König und Ass). Jeder setzt am Anfang einen Dollar in den Topf. Dann zieht zuerst John eine Karte und schaut sie sich an. Er hat dann zwei Optionen: aufgeben (*fold*) – dann bekommt Oskar den Topf, oder erhöhen (*raise*). Im zweiten Fall hat Oskar zwei Möglichkeiten: entweder er gibt seinerseits auf (und John bekommt den Topf) oder ebenfalls einen Dollar einzahlen und dafür die andere Karte nehmen. Im letzten Fall werden dann die Karten verglichen *call* und der mit der höheren Karte gewinnt den ganzen Topf. Natürlich wird John in jedem Fall erhöhen, wenn er das Ass gezogen hat. Bei einem König hat er aber die beiden Strategien “*bluff*” oder “*no bluff*” zur Verfügung (für die er sich auch schon vor dem Ziehen der Karte entscheiden kann). Für Oskar gibt es immer die beiden Strategien “*call*” und “*no call*”. Stelle diese Situation als spieltheoretisches Modell dar und berechne die optimalen (gemischten) Strategien von John und Oskar unter der Annahme, dass beide Spieler rational handeln und ihren Gewinn maximieren (bzw. Verlust minimieren) wollen (d.h., berechne das Nash Gleichgewicht).
- Antwort: Auszahlungsmatrix für John:

$$\begin{array}{cc} & \text{call} & \text{no call} \\ \text{bluff} & \left(\begin{array}{cc} 0 & 1 \\ 1/2 & 0 \end{array} \right) & \\ \text{no bluff} & & \end{array} \quad (23)$$

Sei nun x die Wahrscheinlichkeit, dass John blufft und y die Wahrscheinlichkeit, dass Oskar called. Dann ist der erwartete *payoff* für John:

$$x(1 - y) + (1 - x)y/2 = x + y/2 - (3/2)xy. \quad (24)$$

Oskar hat denselben *payoff*, nur negativ. Für $x > 1/3$ sollte Oskar immer “callen”, für $x < 1/3$ dagegen nie. Für $x = 1/3$ ist die beste Antwort von Oskar beliebig. Andersrum die bluff-Rate von John genau dann beliebig, wenn Oskar mit Wahrscheinlichkeit $y = 2/3$ called. Es ist leicht zu sehen, dass dies das (eindeutige) Nash Gleichgewicht darstellt.

Wiederholte Spielrunden

Kehren wir zur Situation des Gefangenendilemmas zurück. Es gibt die beiden Strategien “*cooperate*” (C) und “*defect*” (D), identisch für beide Spieler. Eine allgemeine

Auszahlungsmatrix hat die Form

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \left(\begin{array}{cc} R & S \\ T & P \end{array} \right) \end{array} \quad (25)$$

mit der Relation $T > R > P > S$: die “*temptation*” ist größer als der “*reward*” für beidseitige Kooperation und der Lohn für “*punishment*” immer noch besser als der “*sucker’s reward*”. Außerdem soll gelten $R > (T + S)/2$, damit der *payoff* in der Summe über beide Spieler bei Kooperation maximal ist. Die entscheidende Frage ist: Wie kann in einer solchen Situation Kooperation evolvieren? Wir hatten schon gesehen, dass dies bei einem einmaligen Spiel sicher unmöglich ist, da dann D die dominante Strategie ist (und damit das einzige ESS). Eine sehr interessante (und realistische) Situation entsteht aber, wenn man sich mehr als einmal trifft im Leben. Dazu nehmen wir an:

1. Das Spiel geht (jeweils zwischen zwei Spielern) über mehrere Runden.
2. Abhängig vom Verhalten des Gegenspielers in der Vorrunde kann jeder Spieler sein Verhalten anpassen.

Mit diesen Regeln explodiert die Anzahl der möglichen Startegien. Im Prinzip kann die Entscheidung in der k -ten Spielrunde sowohl von der Rundenzahl als auch vom Spielgeschehen in den bisherigen $k - 1$ Runden abhängen. In jeder Runde können 4 verschiedene Kombinationen gespielt werden: CC , CD , DC und DD . Damit sind $2^4 = 16$ Strategien möglich, die nur die letzte Runde berücksichtigen, $2^{16} = 56536$ Strategien, die die letzten beiden Runden berücksichtigen und allgemein 2^{4^m} Strategien, die die letzten m Runden berücksichtigen. Daneben gibt es noch stochastische Strategien. Einige bekannte Strategien sind:

- *Always cooperate* (ALLC).
- *Always defect* (ALLD).
- *Grim* (GRIM): Kooperiert zunächst, schaltet aber auf *defect* um (und bleibt dabei), sobald der Gegner das erste Mal *defect* gespielt hat.
- *Tit for tat* (TFT): Beginnt mit C (Variante: mit D) und spielt dann immer genau das, was der Gegner in der letzten Runde gespielt hat.
- *Win stay, loose shift* (WSLS): Beginnt mit C (Variante: mit D) und behält die Strategie (C oder D) genau dann bei, wenn in der letzten Runde ein großer *payoff* erzielt wurde (T oder R). Wechselt die Strategie anderenfalls.
- *Random* (RAN): Feste Wahrscheinlichkeit für *cooperate*, jede Runde wird unabhängig mit dieser Wahrscheinlichkeit gespielt.
- Für alle diese Strategien kann man Abweichungen in einzelnen Runden festlegen, z.B. in der letzten Runde ...

Spiel im Kurs über 4 Runden (feste Rundenzahl) mit 4 verschiedenen Partnern:

$$\begin{array}{cc} & C & D \\ \begin{array}{c} C \\ D \end{array} & \left(\begin{array}{cc} R = 3 & S = 0 \\ T = 5 & P = 1 \end{array} \right) \end{array} \quad (26)$$

Die Werte der Spielmatrix entsprechen derjenigen der berühmten *Axelrod Tournaments* (siehe unten).

Welche mathematischen Aussagen kann man über eine beste Strategie machen? Das direkte Aufeinandertreffen zweier Strategien lässt sich leicht behandeln. Zum Beispiel ist die Auszahlungsmatrix von GRIM gegen ALLD bei m Spielrunden:

$$\begin{array}{cc} & \text{GRIM} & \text{ALLD} \\ \begin{array}{c} \text{GRIM} \\ \text{ALLD} \end{array} & \left(\begin{array}{cc} mR & S + (m-1)P \\ T + (m-1)P & mP \end{array} \right) \end{array} \quad (27)$$

Daraus sieht man, dass für $mR > T + (m-1)P$ GRIM ein striktes Nash Gleichgewicht ist. Nach m aufgelöst erhalten wir die Relation:

$$m > \frac{T - P}{R - P}. \quad (28)$$

Für die Werte der Axelrod Spielmatrix erhalten wir $m > 2$. Allerdings ist ALLD auch immer ein Nash Gleichgewicht. In der Replikatorodynamik gilt für die Frequenz x der GRIM Strategien in der Population: Für $x > \hat{x} = (P - S)/(mR - (m-2)P - S - T)$ steigt GRIM in der Frequenz an und setzt sich durch, sonst setzt sich ALLD durch und fixiert in der Population ($\hat{x} = 1/(2m-4)$ im Beispiel mit den Axelrod-Werten). Wir sehen: Bei vielen Spielrunden genügt schon ein sehr kleiner anfänglicher Anteil an GRIM Strategien, damit diese sich durchsetzen - eine Art von Kooperation entsteht. Man sieht leicht, dass TFT gegen ALLD genau die gleiche Auszahlungsmatrix hat. TFT und GRIM sind gegeneinander neutral.

Frage: Welche Strategie setzt sich langfristig durch, wenn man alle möglichen Strategien mit k Runden zulässt?

Ein Problem bei festen Rundenzahlen ist, dass man in der letzten Runde immer "straflos" defect spielen kann. Es ist leicht zu sehen, dass entsprechend abgewandelte Strategien (TFT-D und GRIM-D) die Originalstrategien dominieren. Sie werden sich also durchsetzen. Dann aber kann man dasselbe Spiel mit der vorletzten Runde spielen, und so weiter. Am Ende landen wir wieder bei ALLD an, welches das einzige Nash Gleichgewicht in diesem erweiterten Strategienraum ist. Da es ein striktes Nash-Gleichgewicht ist ist es auch ein ESS. (In Experimenten mit realen Menschen wird dieser Rückschluss jenseits der letzten Runde in der Regel nicht gemacht.) Dem Problem mit festen Rundenzahlen kann man aber leicht dadurch begegnen, dass man jeweils mit Wahrscheinlichkeit w eine weitere Runde spielt. Die Anzahl der Runden ist dann geometrisch verteilt mit Mittelwert $\bar{m} = 1/(1-w)$. Da der *payoff* linear in m ist, gelten alle Rechnungen von oben dann mit diesem Mittelwert (Axelrod: $w = 0.99654$, $\bar{m} \approx 289$). In diesem Fall ist die langfristige Dynamik ein offenes Problem.

Axelrod's tournaments

Robert Axelrod 1978: Wettbewerb zwischen Computercodes im wiederholten Gefangenendilemma "jeder gegen jeden - inklusive gegen sich selbst". 14 Programme im ersten Wettbewerb, 63 Programme bei der Wiederholung im Jahr 1980. Dabei waren auch viele ausgefeilte Programme, z.B. vom Typ "TESTER", die versuchen, zunächst auszutesten, wie der Gegner auf verschiedene Verhalten reagiert, um das eigene Verhalten dann zu optimieren. (Wenn der Gegner milde reagiert, wird öfter *defect* gewählt, etc). Andere Strategien basieren beispielsweise auf TFT oder GRIM, versuchen aber den Gegner gelegentlich leicht zu übervorteilen, um einen direkten Sieg herauszuholen. Daneben wurde die zufallsbasierte Strategie RANDOM vom Veranstalter selbst hinzugefügt.

- Ein Axelrod-Simulator auf dem Web:
<http://www2.econ.iastate.edu/tesfatsi/demos/axelrod/axelrodt.htm>
- Vereinfachtes Spiel in der Vorlesung: Gruppen zu je 4 Studenten spielen über 6 Runden jeder gegen jeden. Nur deterministische Strategien sind zugelassen (diese aber beliebig). Die 8 besten der Vorlesung spielen in zwei 4-er Gruppen eine zweite 6-er Runde, die 4 besten daraus das Finale. Die Strategie darf zwischen den Runden nicht gewechselt werden.

Gewinner der Original-Turniere war jedes Mal Anatol Rapaport mit der einfachen Strategie TFT. Dies obwohl spätestens beim zweiten Mal alle explizit versuchen, eine Strategie zu entwerfen, die auch TFT schlägt. Robert Axelrod schrieb 1984 ein Buch über die Resultate der *tournaments*, in dem er die Vorzüge von TFT beschreibt (*The Evolution of Cooperation*):

- TFT gewinnt in keinem Spiel mehr als der Gegner – aber es ist auch höchstens marginal schlechter.
- TFT verbraucht keine Energie damit, eine andere Strategie zuerst erkennen zu wollen.
- TFT bestraft *defectors* zuverlässig und kooperiert mit *cooperators* – einschließlich mit sich selbst. Im Gegensatz zu GRIM ist sie prinzipiell versöhnungsbereit. In der Summe der Begegnungen fährt TFT damit sehr gut.
- Die Überlegenheit von TFT gilt aber nur im Spiel gegen verschiedene andere Strategien, die auf Gewinnmaximierung ausgelegt sind. Gegen RANDOM und andere Zufallsstrategien spielt sie sehr schlecht.

Spiele mit Fehlern

Bisher haben wir angenommen, dass beide Spieler in ihren Handlungen ohne Fehler sind. Das heißt insbesondere, dass sie die Handlungen des Gegenspielers richtig interpretieren und dass sie volle Kontrolle über die eigenen Handlungen haben. In der Praxis ist das meist nicht realistisch. Man sollte deshalb bei der Beurteilung von Spielstrategien auch die Möglichkeit von Fehlern berücksichtigen.

- Frage: Nimm an, dass zwei Spieler mit einer geringen *error rate* ϵ TFT gegeneinander spielen. Wie hoch ist dann der durchschnittliche *payoff* pro Runde im Grenzfall sehr vieler Runden?
- Antwort: Markov-Kette mit 3 Zuständen: C (*cooperate*), A (*alternate*), D (*defect*), mit Übergangswahrscheinlichkeit 2ϵ von C nach A und von D nach A und je ϵ von A nach C und D. Damit ist die Übergangsmatrix

$$\begin{pmatrix} (1-\epsilon)^2 & 2\epsilon(1-\epsilon) & \epsilon^2 \\ \epsilon(1-\epsilon) & (1-\epsilon)^2 + \epsilon^2 & \epsilon(1-\epsilon) \\ \epsilon^2 & 2\epsilon(1-\epsilon) & (1-\epsilon)^2 \end{pmatrix} \quad (29)$$

mit führendem linken Eigenvektor $(1/4, 1/2, 1/4)$. Im Gleichgewicht ist der Prozess deshalb die Hälfte der Zeit im Zustand A und je ein Viertel im Zustand C und D. Man bekommt somit einen erwarteten *payoff* von $(R + T + P + S)/4$ pro Runde. Dies ist genau der selbe *payoff* wie bei zwei reinen RANDOM Strategien ($\epsilon = 0.5$).

- Wir sehen: TFT ist anfällig gegen Irrtümer – vor allem auch im Spiel gegen sich selbst. In Wettbewerben mit Fehlerraten führt dies zu einem Vorteil “vergebender” Strategien, die mit einer gewissen Wahrscheinlichkeit nicht bestrafen.

Eine weitere Schwäche von TFT zeigt sich in evolutionären Spielen, wenn Strategien sich mit der Zeit (durch Mutationen) ändern können.

1. Frage: Stelle die *payoff* Matrix von TFT gegen ALLC auf. Wie sieht es mit evolutionärer Stabilität aus?
2. In evolutionären Szenarien ist TFT nur neutral gegen voll kooperative Strategien wie ALLC. Wenn aber zahlreiche ALLC Typen existieren, kann die Population wieder von ALLD übernommen werden.

Reaktive Strategien

Um analytisch weitere Fortschritte machen zu können, müssen wir den Raum der möglichen Strategien eingrenzen. Ein geeignete Wahl sind Strategien mit einer beschränkten Erinnerung, die das Verhalten nur vom Verhalten des Gegenspielers in der vorhergehenden Runde abhängig machen. Deterministisch gibt es dann nur vier Strategien: ALLC, ALLD, TFT und eine paradoxe Strategie PARA, die immer das Gegenteil des vorhergehenden Zuges des Gegners spielt. Probabilistisch können wir den Raum erweitern. Eine Strategie ist dann durch das Tupel (p, q) gekennzeichnet, wobei p die Wahrscheinlichkeit angibt zu kooperieren, wenn der Gegner in der letzten Runde kooperiert hat und q die Kooperations-Wahrscheinlichkeit wenn er nicht kooperiert hat. Man kann diese Strategien auf den Einheitsquadrat darstellen.

Mathematisch kann man nun die Resultate der einzelnen Spielrunden zwischen Spielern mit Strategien (p_1, q_1) und (p_2, q_2) als Markov-Kette darstellen. Die Übergangsmatrix

\mathbf{M} ist leicht konstruiert (m_{ij} Übergang von Zustand i in Zustand j , wobei $i, j \in \{CC, CD, DC, DD\}$, siehe Nowak S. 83),

$$\mathbf{M} = \begin{pmatrix} p_1 p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ q_1 p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ p_1 q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ q_1 q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{pmatrix}. \quad (30)$$

Definiere dann \mathbf{x}_t als (auf 1 normierten) Wahrscheinlichkeitsvektor für den Zustand des Spiels in der t -ten Runde. Dann ist

$$\mathbf{x}_{t+1} = \mathbf{x}_t \mathbf{M}. \quad (31)$$

Da alle Zeilen von \mathbf{M} sich zu 1 addieren ist jeder Vektor \mathbf{x}_t wieder auf 1 normiert. Offensichtlich ist 1 damit auch Eigenwert von \mathbf{M} . Nehmen wir nun an, dass \mathbf{M} *primitiv* ist, d.h. für eine positive Zahl k sind alle Einträge von \mathbf{M} positiv. Dann gilt der Satz von *Perron-Frobenius*: es gibt genau einen maximalen Eigenwert $\lambda = 1$ mit zugehörigem linken Eigenvektor \mathbf{v} . Dieser hat nur positive Einträge, die die Gleichgewichtsverteilung des Prozesses angeben. Man findet:

$$\mathbf{v} = (s_1 s_2, s_1(1-s_2), (1-s_1)s_2, (1-s_1)(1-s_2)) \quad (32)$$

mit

$$s_1 = \frac{q_2(p_1 - q_1) + q_1}{1 - (p_1 - q_1)(p_2 - q_2)} \quad ; \quad s_2 = \frac{q_1(p_2 - q_2) + q_2}{1 - (p_2 - q_2)(p_1 - q_1)}. \quad (33)$$

s_1 bzw. s_2 ist die Wahrscheinlichkeit, dass Spieler 1 bzw. 2 im Gleichgewicht kooperiert. Die Gleichgewichtsverteilung in der Form (32) gilt immer dann, wenn es nur einen einzigen Eigenvektor von \mathbf{M} zum Eigenwert $\lambda = 1$ gibt. Dies ist fast immer der Fall, außer im Spiel von TFT (1, 0) oder von PARA (0, 1) gegen sich selbst. In diesen speziellen Fällen hängt die Gleichgewichtsverteilung vom Anfangszustand (Spiel in der ersten Runde) ab. Der erwartete *payoff* pro Runde für den Spieler mit Strategie 1 gegen Strategie 2 berechnet sich dann leicht zu

$$E_{12} = R s_1 s_2 + S s_1(1-s_2) + T(1-s_1)s_2 + P(1-s_1)(1-s_2). \quad (34)$$

Diese Werte kann man nun allgemein verwenden, um eine *payoff* Matrix \mathbf{E} mit Einträgen E_{ij} für das Spiel zwischen n beliebigen reaktiven Strategien (p_i, q_i) , $1 \leq i \leq n$, im Limes großer Rundenzahlen zu konstruieren.

Für ein evolutionäres Szenario beginnen wir mit n zufällig bestimmten Strategien, die wir gemäß der Replikatorgleichung mit der Auszahlungsmatrix \mathbf{E} evolvieren lassen. In Simulationen zeigt sich meist ein großer anfänglicher Vorteil von ALLD-nahen Strategien (p, q nahe 0, 0). Aber dann, aufgrund der hohen Rundenzahlen (für jede Generation des evolutionären Spiels), können sich TFT-nahe Strategien behaupten und ausbreiten. Dies ist aber noch nicht das Ende: sobald ALLD sehr selten ist, breitet sich eine Variante von TFT aus, die Fehler verzeiht: GTFT *generous tit for tat*. Für die Axelrod Werte ergibt sich im Grenzfall eine Strategie mit $(p, q) = (1, 1/3)$. Man kann zeigen, dass dies

die Strategie mit dem maximalen *payoff* im Spiel gegen sich selbst ist (mit Fehlern), die noch gegen Invasion von eigennützigeren Strategien mit kleinerem p und/oder q stabil ist.

Man kann die Analyse nun weitertreiben und fragen, was die beste Strategie ist, die nicht nur den letzten Zug des Gegeners, sondern auch den letzten eigenen Zug berücksichtigt. Man erhält dann einen 4-dim. Einheitskubus als Raum aller Strategien. Als sehr erfolgreiche Strategie in einer bereits gut kooperierenden Population mit Fehlerraten stellt sich WSLS heraus. Obwohl die Strategie weniger gut gegen ALLD ist als TFT, ist sie doch stabil und hat zwei Vorteile:

- WSLS kann bei Fehlern im Spiel gegen sich selbst deterministisch korrigieren (besser als die stochastische Korrektur von GTFT).
- WSLS (mit Fehlern) hat einen hohen *payoff* gegen ALLC. Dies ist vor allem dann ein Vorteil, wenn durch neutrale Evolution (Drift) ALLC sich in einer sehr gut kooperierenden Population ausbreiten kann. Man bekommt eine Dynamik wie auf S. 91 von Nowak dargestellt.

Fragen:

- Betrachte GTFT mit $p = 1$ und allgemeinem q gegen ALLD im unendlich oft iterierten Gefangenendilemma. Wie sieht die Auszahlungsmatrix für die mittlere Auszahlung pro Runde aus? Wann ist GTFT noch ein Nash-Gleichgewicht?
- Betrachte die deterministische Strategie WSLS gegen ALLD in einem Spiel über k Runden. Bestimme wieder die Auszahlungsmatrix. Für welche Werte von R , T , S , P ist WSLS noch ein Nash-Gleichgewicht?

Evolution in endlichen Populationen

Bisher haben wir für die evolutionäre Spieltheorie angenommen, dass die mittlere Anzahl der Nachkommen einer bestimmten Strategie genau dem mittleren *payoff* in der Population entspricht, wenn alle gegen alle spielen. In jeder realen Situation wird aber die tatsächliche Anzahl an Nachkommen eines einzelnen Individuums eine Zufallsvariable sein, die um den entsprechenden Mittelwert herum schwankt. In unendlich großen Populationen mitteln sich diese Schwankungen heraus, wenn wir nicht die Nachkommen eines einzelnen Individuums betrachten, sondern die Frequenz einer Strategie in der Population. In endlichen Populationen ist dies aber nicht der Fall: hier hängt die Änderung einer Allelfrequenz (hier: Frequenz einer Strategie) sowohl von der Fitness ab, als auch von den Zufallsschwankungen. In der Biologie bezeichnet man diese Zufallsschwankungen als *genetische Drift*.

Wir betrachten eine Population mit einer festen Größe N . Individuen treten in zwei Arten auf: A und a , die verschiedene Genotypen (Allele, Strategien) repräsentieren. Man kann dann den Zustand der Population zur Zeit t einfach durch die Anzahl der A Individuen angeben (bzw. durch ihre Frequenz). Man hat also $N + 1$ Zustände von 0

bis N . Wie beschreibt man die Evolution in solch einer endlichen Population als mathematisches Modell? Ein Standardmodell in diskreten Generationen ist das Wright-Fisher Modell. Für die Spieltheorie geeigneter ist ein Modell mit überlappenden Generationen, das sogenannte Moran Modell.

Moran Modell

Ein Modell für Drift mit überlappenden Generationen ist das Moran Modell. Die Idee ist, dass man einzelne Geburten und Tode beschreibt und beide Prozesse koppelt, damit die Populationsgröße erhalten bleibt. Dazu werden pro Zeiteinheit zwei Individuen ausgewählt (ggf. auch das gleiche zweimal): eines zur Reproduktion, das andere zum Sterben. Wir können uns z.B. vorstellen, dass durch einen Tod ein Platz für ein neues Individuum frei geworden ist. Wir definieren den Zustandsvektor $\mathbf{x}(t) = (x_0(t), \dots, x_N(t))$, wobei $x_i(t)$ die Wahrscheinlichkeit angibt, dass zum Zeitpunkt t genau i Individuen die Strategie (das Allel) A tragen. Die Dynamik von $\mathbf{x}(t)$ lässt sich leicht als Markov-Prozess darstellen, $\mathbf{x}(t+1) = \mathbf{x}(t)\mathbf{P}$, mit Übergangsmatrix \mathbf{P} .

- Wie lauten die Einträge $p_{k,l}$ von \mathbf{P} im neutralen Fall ohne Mutation und Selektion?

$$p_{k,k+1} = \frac{k(N-k)}{N^2} \quad (35)$$

$$p_{k,k-1} = \frac{k(N-k)}{N^2} = p_{k,k+1} \quad (36)$$

$$p_{k,k} = 1 - p_{k,k+1} - p_{k,k-1} \quad (37)$$

Alle anderen Einträge sind Null: die Matrix \mathbf{P} ist tridiagonal. Einen solchen Prozess bezeichnet man allgemein auch als Geburts-Todesprozess.

Elementare Eigenschaften:

1. Ohne Berücksichtigung von Neumutationen gibt es zwei absorbierende Zustände (Fixation) bei $k=0$ und $k=N$. Langfristig wird der Evolutionsprozess in einem dieser absorbierenden Zustände enden.
2. Sei $k(t)$ die Anzahl der Individuen mit Allel A zur Zeit t und $E[k(t)] = \sum_{i=0}^N x_i(t)i$ ihr Erwartungswert. Im neutralen Fall (keine Selektion) haben wir $p_{k,k-1} = p_{k,k+1}$. Damit berechnet man sehr leicht, dass gilt $E[k(t+1)] = E[k(t)]$. Man sagt auch, der Prozess ist erwartungserhaltend (ein Martingal).
3. Die Wahrscheinlichkeit u der Fixation bei N (und nicht bei 0) für einen bei $k(0) = k_0$ gestarteten neutralen Prozess ergibt sich direkt aus der Martingaleigenschaft. Mit $k_0 = E[k(0)] = \lim_{t \rightarrow \infty} [E[k(t)]] = u \cdot N + (1-u) \cdot 0$ folgt

$$u[k(0) = k_0] = k_0/N.$$

Selektion kann im Moran Modell entweder über die Reproduktions- oder die Todesraten eingeführt werden. Über die Reproduktion können wir etwa definieren: Sei r die Fitness eines Mutanten. Dann ist

$$r = \frac{Pr[\text{repro. mutant}]}{Pr[\text{repro. wt}]}$$

die Wahrscheinlichkeit, dass ein Mutant zur Reproduktion ausgewählt wird, relativ zur Wahrscheinlichkeit, dass ein Wildtyp ausgewählt wird. Damit ist insgesamt die Wahrscheinlichkeit, dass ein Mutanten A Allel ausgewählt wird wenn i solche Allele vorhanden sind:

$$\frac{r \cdot i}{r \cdot i + N - i} \quad (38)$$

Entsprechend ist die Wahrscheinlichkeit, dass ein Wildtyp a Allel ausgewählt wird

$$\frac{N - i}{r \cdot i + N - i}. \quad (39)$$

Die Todesraten bleiben für beide Typen wie im neutralen Fall. Damit ist also

$$p_{i,i+1} = \frac{r \cdot i(N - i)}{(r \cdot i + N - i)N} \quad (40)$$

$$p_{i,i-1} = \frac{i(N - i)}{(r \cdot i + N - i)N} \quad (41)$$

Mit Selektion ist der Geburts-Todesprozess nicht mehr erwartungserhaltend. Wie berechnet man jetzt die Fixationswahrscheinlichkeiten? Sei u_i die Fixationswahrscheinlichkeit bei N für einen Prozess, der bei $k(0) = i$ gestartet wird. Wir haben:

$$u_0 = 0 \quad (42)$$

$$u_i = p_{i,i+1}u_{i+1} + p_{i,i-1}u_{i-1} + p_{i,i}u_i \quad (43)$$

$$u_N = 1 \quad (44)$$

Damit ist insgesamt

$$\mathbf{u} = \mathbf{P}\mathbf{u} \quad (45)$$

Der Vektor \mathbf{u} ist also ein (Rechts-) Eigenvektor von \mathbf{P} zum Eigenwert 1. Aufgrund der tridiagonalen Form der Matrix können wir \mathbf{u} für einen beliebigen Geburts-Todesprozess mit tridiagonaler Übergangsmatrix berechnen. Dazu definieren wir $y_i = u_i - u_{i-1}$. Dann ist

$$\sum_{i=1}^N y_i = u_N - u_0 = 1. \quad (46)$$

Ausserdem erhalten wir mit $p_{i,i} = 1 - p_{i,i+1} - p_{i,i-1}$ aus (43)

$$y_{i+1} = \frac{p_{i,i-1}}{p_{i,i+1}}y_i =: \gamma_i y_i$$

und wegen $y_1 = u_1$ auch

$$y_k = u_1 \prod_{i=1}^{k-1} \gamma_i$$

und damit auch

$$u_1 = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^k \gamma_i}$$

und

$$u_j = \sum_{k=1}^j y_k = u_1 \left(1 + \sum_{k=1}^{j-1} \prod_{i=1}^k \gamma_i \right) = \frac{1 + \sum_{k=1}^{j-1} \prod_{i=1}^k \gamma_i}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^k \gamma_i}.$$

- Im neutralen Fall ist $p_{i,i-1} = p_{i,i+1}$ und damit $\gamma_i = 1$. Wir erhalten sofort

$$u_j = \frac{j}{N}$$

wie erwartet.

- Mit Selektion haben wir

$$\gamma_i = \frac{p_{i,i-1}}{p_{i,i+1}} = \frac{1}{r}.$$

Wir erhalten also

$$u_i = \frac{\sum_{j=0}^{i-1} r^{-j}}{\sum_{j=0}^{N-1} r^{-j}} = \frac{(1 - r^{-i})/(1 - r^{-1})}{(1 - r^{-N})/(1 - r^{-1})} = \frac{1 - 1/r^i}{1 - 1/r^N}. \quad (47)$$

Die Fixationswahrscheinlichkeit eines einzelnen A Mutanten ist also:

$$\rho_A = u_1 = \frac{1 - 1/r}{1 - 1/r^N} \approx (r - 1)/r \quad (48)$$

für $r \approx 1$.

- Umgekehrt ist die Fixationswahrscheinlichkeit für einen einzelnen Wildtypen in der Mutantenpopulation

$$\rho_a = 1 - u_{N-1} = \frac{1 - r}{1 - r^N} \approx (r - 1)/r^N. \quad (49)$$

Wir haben immer

$$\frac{\rho_a}{\rho_A} = r^{1-N}.$$

Biologisch sind Fixationswahrscheinlichkeiten selbst für stark positive Mutationen typischerweise sehr klein. Für schädliche Mutationen werden sie exponentiell klein. Für eine Populationsgröße $N = 100$ erhalten wir die folgenden Fixationswahrscheinlichkeiten für

eine einzelne Neumutation und die Anzahl $\#_{\rho=1/2}$ der benötigten A Neumutationen, dass wir mit vorgegebener Wahrscheinlichkeit $\rho = 1/2$ Fixation von A bekommen:

Fitness	Fixationswahrsch.	$\#_{\rho=1/2} = -\log(2)/\log(1-\rho)$
$r = 2$ (+100%)	$\rho = 0.5$	1
$r = 1.1$ (+10%)	$\rho = 0.09$	7
$r = 1.01$ (+1%)	$\rho = 0.016$	44
$r = 1$ (neutral)	$\rho = 1/N = 0.01$	69
$r = 0.99$ (-1%)	$\rho = 0.0058$	119
$r = 0.9$ (-10%)	$\rho = 0.000003$	234861

Evolutionäre Raten

Wir betrachten eine Population der Größe N . Die Mutationsrate pro Individuum sei u . Dann treten in der Population pro Generation (bzw. pro Zeiteinheit) Nu Mutationen auf. Die Zeit, die wir warten müssen bis zum Auftreten der nächsten Mutation ist exponentiell verteilt mit Parameter $\mu = Nu$. Falls eine neue Mutation auftritt, besteht die Wahrscheinlichkeit ρ , dass sie fixiert. Wir können also sagen: die Rate der auftretenden "erfolgreichen" Mutationen ist

$$R = Nu\rho \quad (50)$$

Für die neutrale Evolution ist

$$R = Nu \frac{1}{N} = u \quad (51)$$

Man sagt auch: Die neutrale Substitutionsrate entspricht der neutralen Mutationsrate. Dies ist ein wichtiges Ergebnis: Wenn die Mutationsraten über die Zeit konstant bleiben, dann sind auch die Substitutionsraten konstant (unabhängig) von Änderungen in den Populationsgrößen. Dies ist die Basis für die sogenannte *molekulare Uhr*, die es ermöglicht aus dem Vergleich von Sequenzdaten phylogenetische Datierungen abzuleiten.

Wir vergleichen homologe DNA Abschnitte von Individuen in zwei verschiedenen Spezies (z.B. Mensch - Schimpanse). Nehmen wir an, dass vor genau T Generationen der jüngste gemeinsame Vorfahre von beiden lebte. Dann trennen beide Individuen $2T$ Generationen. Die erwartete Anzahl an Substitutionen in dieser Zeit ist dann Poisson verteilt mit Parameter $\lambda = 2TR$

- Poisson Verteilung mit Parameter λ :

$$Pr[X = k] = \exp[-\lambda] \cdot \frac{\lambda^k}{k!}$$

Mittelwert und Varianz:

$$E[X] = \text{Var}[X] = \lambda.$$

- Dies können wir verwenden, um aus der Anzahl der Substitutionen \hat{X} die Zeit seit der Trennung zu schätzen.

$$E[X] = 2TR \quad \Rightarrow \quad \hat{T} = \frac{\hat{X}}{2R}. \quad (52)$$

Alternativ kann man die Mutationsrate schätzen, wenn die Zeit seit der Trennung bekannt ist (wie wäre dann die Gleichung?)

- Beispiel: Wir vergleichen eine (homologe) Sequenz der Länge 8000 Basenpaare eines Menschen und eines Schimpansen. Wir finden 133 Substitutionen in dieser Sequenz. Bei einer Mutationsrate $u = 4 \cdot 10^{-8}$ pro Generation und Basenpaar und einer angenommenen mittleren Generationszeit von 25 Jahren, wie groß ist dann die geschätzte Zeit bis zum gemeinsamen Vorfahren von Mensch und Schimpanse? [$\hat{T} = 25 \cdot 133 / (2 \cdot 8000 \cdot 4 \cdot 10^{-8}) = 5.2 \cdot 10^6$ Jahre].
- Aufgabe: Wenn auf einer einzigen Basenposition mehrere Substitutionen stattfinden, dann ist dies an den resultierenden Sequenzen nicht mehr erkennbar. Betrachte zwei Spezies mit einer Mutationsrate $u = 4 \cdot 10^{-8}$ pro Basenpaar und Generation. Ab welcher Distanz (in Generationen) zwischen zwei Spezies ist in $> 5\%$ aller sichtbaren Substitutionen eine Mehrfachsubstitution zu erwarten (und damit die Schätzung aus der molekularen Uhr schwieriger)?

Die molekulare Uhr ist nicht immer erfüllt: auch Mutationsraten können sich über evolutionäre Zeiträume ändern. Eine Möglichkeit, die zu bestimmen ist der *relative rates test*. [Was für Daten braucht man, um so etwas testen zu können?]

Baum mit (mindestens) drei Arten. Idee: Vergleiche die Raten der beiden gleich langen Äste. Zum Beispiel: Mensch - Altweltaffen (Rhesus Macaque) - Neuweltaffen (Kapuziener) als *outgroup*.

$$X_{OA} = \frac{X_{AB} + X_{AC} - X_{BC}}{2} \quad ; \quad X_{OB} = \frac{X_{AB} + X_{BC} - X_{AC}}{2}$$

Wenn wir annehmen, dass Mutationen auf den Ästen OA und OB mit der gleichen Rate auftreten, dann sollte jede sichtbare Substitution $X_{ges} = X_{OA} + X_{OB}$ mit Wahrscheinlichkeit $1/2$ auf einem der beiden Äste liegen. Dies führt auf eine Binomialverteilung und kann entsprechend getestet werden.

$$Pr[X = k] = p^k (1 - p)^{n-k} \binom{n}{k} \quad (53)$$

wobei hier $n = X_{ges}$ und $p = 1/2$ ist. Wir haben dann

$$Pr[X \leq X_{OA}] = 2^{-X_{ges}} \sum_{k=0}^{X_{OA}} \binom{X_{ges}}{k} \quad (54)$$

Für einen zweiseitigen Test ist dann der p -Wert $p = 2Pr[X \leq X_{OA}]$ (muss mit dem Computer bestimmt werden).

- Zahlenbeispiel: Distanz Mensch – Altweltaffe: $k_{AB} = 485$ Substitutionen, Distanz Mensch – Neuweltaffe: $k_{AC} = 1201$, und Distanz Altweltaffe – Neuweltaffe: $k_{BC} = 1288$ Substitutionen. Dann ist $k_{OA} = 199$ und $k_{OB} = 286$. Als p -Wert ergibt sich $p \approx 4.5 \cdot 10^{-5}$. Die molekulare Uhr ist demnach für diese Daten stark gestört.

Spiele in endlichen Populationen

Die klassische (evolutionäre) Spieltheorie ist eine deterministische Theorie. In der Replikatorgleichung werden Zufallsschwankungen durch Drift nicht berücksichtigt: Populationen sind effektiv “unendlich groß”. Bei einer endlichen Anzahl von Begegnungen lassen sich stochastische Effekte aber nicht ignorieren. Es stellt sich die Frage nach einer Erweiterung der Spieltheorie auf endliche Populationen. Bisher hatten wir:

$$\dot{x}_i = x_i(w_i - \bar{w}) \quad (55)$$

mit

$$w_i[\mathbf{x}] = \sum_{j=1}^n a_{ij}x_j \quad \text{und} \quad \bar{w} = \sum_{i,j=1}^n a_{ij}x_i x_j \quad (56)$$

Wir wollen nun die Fitness in Abhängigkeit des *payoffs* etwas allgemeiner definieren:

$$w_i[\mathbf{x}] \rightarrow r_i[\mathbf{x}] := 1 - w + w w_i[\mathbf{x}] \quad (57)$$

Der zusätzliche Parameter w ($0 < w \leq 1$) erlaubt es, die Größe der Fitnessunterschiede (= Stärke der Selektion) zu regulieren. Für das deterministische System ändert dieser Parameter aber nichts entscheidendes an der Dynamik. Um dies zu sehen, verwende r_i in der Replikatorgleichung: Die Konstante $1 - w$ fällt über die mittlere Fitness wieder heraus. Der Faktor w bewirkt im deterministischen Modell nur eine Reskalierung der Zeitskala: mit schwacher Selektion konvergieren wir langsamer zum Gleichgewicht. Dies spielt keine Rolle wenn wir uns nur für die Gleichgewichte (das Langzeit-Verhalten) interessieren.

Um ein alternatives Modell mit Drift zu erhalten, können wir die gleiche Fitness $r_i[\mathbf{x}]$ in einem Moran Modell verwenden. Der Einfachheit halber beschränken wir uns auf zwei Typen (Strategien), $i = 1, 2$ (entsprechend den Allelen a und A im letzten Abschnitt). Die Auszahlungsmatrix ist

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (58)$$

Mit $x_1 = k/N$ und $x_2 = (N - k)/N$ können wir die Fitnesswerte, die sich aus dem Spiel ergeben, dann schreiben als

$$r_1[\mathbf{x}] = r_1[k] = 1 - w + w \frac{a(k-1) + b(N-k)}{N-1}$$

und

$$r_2[\mathbf{x}] = r_2[k] = 1 - w + w \frac{ck + d(N-k-1)}{N-1}$$

und bekommen die Übergangsraten:

$$p_{k,k+1} = \frac{kr_1[k]}{kr_1[k] + (N-k)r_2[k]} \frac{N-k}{N}, \quad (59)$$

$$p_{k,k-1} = \frac{(N-k)r_2[k]}{kr_1[k] + (N-k)r_2[k]} \frac{k}{N}. \quad (60)$$

Dies entspricht einem Selektionsmodell mit frequenzabhängiger Fitness. Um evolutionäre Raten in beiden Richtungen zu bestimmen, interessieren wir uns für die Fixationswahrscheinlichkeiten. Es gilt

$$\frac{p_{k,k-1}}{p_{k,k+1}} = \frac{r_2[k]}{r_1[k]} \approx 1 - w \frac{a(k-1) + b(N-k) - ck - d(N-k-1)}{N-1}$$

wenn wir für schwache Selektion bis zur linearen Ordnung in w entwickeln. Wir erhalten dann für die Fixationswahrscheinlichkeit einer "Neumutation" der Strategie A

$$\rho_A = \left(1 + \sum_{k=1}^{N-1} \prod_{i=1}^k \frac{r_2[i]}{r_1[i]} \right)^{-1} \quad (61)$$

$$\approx \left(1 + \sum_{k=1}^{N-1} \left(1 - w \sum_{i=1}^k \frac{a(i-1) + b(N-i) - ci - d(N-i-1)}{N-1} \right) \right)^{-1} \quad (62)$$

$$= \left(N - \frac{w}{N-1} \sum_{k=1}^{N-1} \sum_{i=1}^k (a(i-1) + b(N-i) - ci - d(N-i-1)) \right)^{-1} \quad (63)$$

$$= \frac{1}{N} \cdot \frac{1}{1 - \frac{w}{6}(a(N-2) + b(2N-1) - c(N+1) - d(2N-4))} \quad (64)$$

wobei wir verwenden

$$\sum_{i=1}^k i = \frac{k(k+1)}{2} \quad \text{und} \quad \sum_{i=1}^k i^2 = \frac{k(k+1)(2k+1)}{6}.$$

Die Selektion fördert genau dann die Fixation, wenn $\rho_A > 1/N$ ist. Dies ist genau dann der Fall wenn

$$a(N-2) + b(2N-1) > c(N+1) + d(2N-4) \quad (65)$$

Für eine sehr kleine Population von $N = 2$ führt das auf die Bedingung $b > c$, wie erwartet. Für eine sehr große Population erhalten wir

$$a + 2b > c + 2d. \quad (66)$$

Interessant ist diese Bedingung für den Fall, dass beide Strategien lokale Gleichgewichte sind (Nash und ESS). Dies ist der Fall, wenn $a > c$ und $b < d$ ist. Es gibt dann ein inneres instabiles Gleichgewicht der Replikatorgleichung, das die Einzugsbereiche der Gleichgewichte trennt. Für großes N ist die Frequenz der Strategie A an diesem Gleichgewicht

$$x^* = \frac{d-b}{a-b-c+d}$$

Mit der Bedingung (65) bekommen wir jetzt gerade

$$x^* < \frac{1}{3} \quad (67)$$

Wenn also der Einzugsbereich von Strategie A mindestens $2/3$ des Frequenzraums umfasst, dann ist die Fixation von A durch Selektion begünstigt.

Wenn $a > c$ und $b > d$ gilt, dann dominiert die Strategie A über die alternative Strategie (zB B). Da dann gilt dass $x^* < 0$ ist, wird die Fixation von A in einer großen Population auch immer von Selektion begünstigt sein. In einer kleinen Population gilt dies aber nicht: wenn $b < c$ ist, dann ist für eine kleine Population B in der Fixation begünstigt!

Evolutionäre Stabilität in endlichen Populationen

Evolutionäre Stabilität heißt, dass ein Zustand gegen das Eindringen von Mutanten (in anfänglich geringer Frequenz) geschützt ist. In einer unendlichen Population ergeben sich mit der Auszahlungsmatrix die folgenden Bedingungen für Strategie B als ESS:

1. $d \geq b$ (Nash)
2. Wenn $d = b$, dann $a < c$.

In einer endlichen Population, können wir ein stochastisches “ESS $_N$ ” definieren als

1.

$$c + d(N - 2) \geq b(N - 1) \tag{68}$$

Also: Der *payoff* eines einzelnen B Mutanten in einer B Population ist größer (oder gleich) wie der eines A Typs (A wird durch Selektion unterdrückt).

2.

$$a(N - 2) + b(2N - 1) < c(N + 1) + d(2N - 4) \tag{69}$$

Dies bedeutet, dass $\rho_A < 1/N$ ist, also, dass Selektion die Substitution von B durch A unterdrückt.

Im Limes sehr kleiner Populationen $N \rightarrow 2$ reduzieren sich die Bedingungen auf $b < c$. Dies ist sehr verschieden von den klassischen ESS Bedingungen. Der Grund ist klar: in einer kleinen Population ist selbst ein einzelner neuer Mutant nicht “infinitesimal selten”. Im Grenzfall $N = 2$ stellt er die halbe Population dar. Da wir den Fall ausgeschlossen haben, dass ein Spieler gegen sich selbst spielt, unterscheiden sich dann die Frequenzen der Strategien der Gegenspieler deutlich von den Frequenzen in der Gesamtpopulation. Für große Populationen ($N \rightarrow \infty$) erhalten wir $d \geq b$ und $x^* > 1/3$. Ein klassisches ESS ist dann eine notwendige Bedingung, aber nicht hinreichend. Der Grund in diesem Fall ist, dass auch ein lokal durch Selektion stabilisiertes Gleichgewicht bezüglich Substitution instabil sein kann. Das ist der Fall, wenn Selektion auf dem größten Teil des Frequenzraums (mehr als $2/3$) das Anwachsen der Mutantenfrequenz fördert: Man schwimmt einen genügend großen Teil des Weges mit dem Strom.

Risikodominanz

Wenn sowohl Strategie A als auch B Nash Gleichgewichte sind, kann man in einer endlichen Population auch fragen, welche Strategie die andere mit größerer Wahrscheinlichkeit ersetzen kann. Wir interessieren uns also dafür, wann $\rho_A > \rho_B$ gilt. Mit

$$\rho_A = \frac{1}{N} \cdot \frac{1}{1 - \frac{w}{6}(a(N-2) + b(2N-1) - c(N+1) - d(2N-4))} \quad (70)$$

$$\rho_B = \frac{1}{N} \cdot \frac{1}{1 - \frac{w}{6}(d(N-2) + c(2N-1) - b(N+1) - a(2N-4))} \quad (71)$$

folgt, dass dies gerade gilt für

$$(N-2)(a-d) > N(c-b). \quad (72)$$

Für große Populationen ist dies erfüllt wenn

$$a + b > c + d$$

Dies entspricht genau der Bedingung, dass A den größeren Einzugsbereich hat, also $x^* < 1/2$. Man bezeichnet diese Eigenschaft auch als Risikodominanz. Für eine kleine Populationsgröße ändert sich die Bedingung aber (in Richtung $b > c$, was bei $N = 2$ erreicht wird).

Anwendung: Tit-for-tat gegen ALLD in endlichen Populationen

Wir betrachten ein Spiel über m Runden unter den üblichen Regeln für das iterierte Gefangenendilemma. Die Auszahlungsmatrix ist

$$\begin{array}{cc} & \begin{array}{cc} \text{TFT} & \text{ALLD} \end{array} \\ \begin{array}{c} \text{TFT} \\ \text{ALLD} \end{array} & \left(\begin{array}{cc} mR & S + (m-1)P \\ T + (m-1)P & mP \end{array} \right) \end{array} \quad (73)$$

Im deterministischen Spiel (Replikatorgleichung) ist ALLD immer strikt Nash und ein ESS. Für TFT erhalten wir in diesem Fall die Bedingung $mR > T + (m-1)P$ für ein striktes Nash Gleichgewicht (und damit auch ein deterministisches ESS). Für $mR = T + (m-1)P$ ist TFT immer noch Nash, aber nicht strikt und auch kein ESS mehr (da $mP > S + (m-1)P$). Nach m aufgelöst erhalten wir die Bedingung für die evolutionäre Stabilität von TFT im Spiel ohne Drift:

$$m > \frac{T-P}{R-P}. \quad (74)$$

Im zweiten Schritt betrachten wir die stochastische Stabilität. Für ALLD prüft man leicht nach, dass (68) immer erfüllt ist. Für die zweite Bedingung (69)

$$a(N-2) + b(2N-1) < c(N+1) + d(2N-4)$$

erhalten wir nach Einsetzen der Werte der Auszahlungsmatrix:

$$m < \frac{T(N+1) + P(N-2) - S(2N-1)}{(R-P)(N-2)}. \quad (75)$$

Für $R = 3$, $T = 5$, $P = 1$, $S = 0$ wird dies

$$m < \frac{6N+3}{2N-4} \quad (76)$$

Für $N = 2$ ist (69) immer erfüllt. Dies ist intuitiv klar: in einer Population von nur zwei Individuen ist es nie von Vorteil, dem anderen zu helfen. Anders gesagt: ein TFT Mutant ist in jedem Paar-Wettbewerb mit einer nicht kooperierenden Strategie unterlegen. Bereits für $N = 3$ aber wird ALLD ab einer minimalen ($m = 11$ für die Axelrod Werte) stochastisch instabil. Für steigendes N nimmt die kritische Rundenzahl immer weiter ab – bis auf $m = 3$ im Limes sehr großer Populationen.

Analoge Rechnungen im stochastischen Modell können wir auch für TFT durchführen. Wir erhalten so die Stabilitätsbedingung

$$m > \frac{T(2N-1) - P(N-2) - S(N+1)}{2(R_P)(N-2)} \quad (77)$$

oder

$$m > \frac{9N-3}{4N-8} \quad (78)$$

im Axelrod-Fall. Für $N = 2$ ist TFT nie stabil. Für $N = 3$ kommen wir mit den Axelrod-Werten auf $m > 6$. Für $N \rightarrow \infty$ ist TFT schon für $N = 3$ stochastisch stabil.

Im Bereich $3 \leq N < \infty$ finden wir jeweils Werte von m , bei denen sowohl ALLD als auch TFT stochastisch stabil sind. Wir können dann weiter fragen, wann TFT risikodominant ist. Aus der Bedingung (72) bekommen wir leicht

$$m > \frac{N(T-S)}{(N-2)(R-P)} \quad (79)$$

bzw.

$$m > \frac{5N}{2N-4} \quad (80)$$

für die Axelrod-Werte.