Operations Research B (SoSe 04)

Joachim Rosenmüller   Jonas Zimmermann

22. Juli 2004

Inhaltsverzeichnis

Einführung

Modellierungsschritte

1) System, z. B.
   (1) Elektrischer Schwingkreis
   (2) Maschinenpark
   (3) Lager
   (4) Oligopol („Tankstellenmarkt“)
   (5) Volkswirtschaft
2) steuerbar:
   (1) Eingangsspannung, Kapazitäten...
   (2) Reparatur / Erneuerung
   (3) Bestellung, Lagerhaltungspolitik
   (4) Preissetzung, Mengenfestsetzung
   (5) Investitionsprogramme, Besteuerung
3) Nutzen, Auszahlungen, „Incentives“ – Anreize
   (1) Empfangsqualität (eines Fernsehers)
   (2) Lebensdauer, Kostenminimierung
   (3) Lagerkosten, Suchkosten, Kundenärger
   (4) Profite / Gewinne
   (5) BIP, Beschäftigungsquote, Preisstabilität

Annahmen über Systeme

1) System kann beobachtet werden (Zustände, relevante Parameter)
2) volle / eingeschränkte Informationen für Teilnehmer, Kontrolleure, Spieler
3) Kenntnis des Einflusses einer Aktion / mehrerer Aktionen auf die Zustandsänderung
4) Kenntnis von „Nutzenfunktionen“ (rewards, payoffs) abhängig von temporären Aktionen und globalen Strategien:
   - intermediäre Auszahlungen bei Aktionen
   - Endauszahlungen
   - globale Auszahlung ist deren Summe

Erste Aufgabe: Formulierung eines einfachen dynamischen Modells, Optimierung (für einen Entscheidungsträger)

1 Das Prinzip der Dynamischen Optimierung

1.1 Rückwärtsinduktion auf einem Graphen

Wir beginnen mit einer

Definition 1.1.1. *Als* Zustandsraum *interpretiert man eine endliche Menge*

$$\mathfrak{X} = \{\xi, \eta, \ldots, \kappa\}.$$

Ihre Elemente heißen Zustände *oder* Knoten. *Eine* binäre Relation *auf $\mathfrak{X}$ ist eine Teilmenge $\prec \subseteq \mathfrak{X} \times \mathfrak{X}$, wir schreiben $\xi \prec \eta$ statt $(\xi, \eta) \in \prec$ und sagen / interpretieren: „ξ kommt vor η“ oder „η folgt auf ξ“.*
Das Paar $(\mathfrak{X}, \prec)$ heißt Graph.

Wir verlangen bestimmte Eigenschaften von $(\mathfrak{X}, \prec)$, z. B.:
$\prec$ heißt *asymmetrisch*, falls aus $\xi \prec \eta$ stets folgt, daß $\eta \prec \xi$ falsch ist. Insbesondere: $\xi \not\prec \xi$!

Definition 1.1.2. *Eine Folge $x = (x_0, x_1, \ldots, x_T)$ mit $x_t \in \mathfrak{X}$ $\forall t \in \{0, \ldots, T\}$ heißt* Pfad *(der x_0 und x_T verbindet), wenn $x_t \prec x_{t+1} \forall t \in \{0, \ldots, T-1\}$.*
Ein Graph $(\mathfrak{X}, \prec)$ induziert eine weitere Relation $\overset{\mathrm{T}}{\prec}$ (die transitive Hülle *von $\prec$); diese ist definiert vermöge*

$$\xi \overset{\mathrm{T}}{\prec} \eta \;:\Longleftrightarrow\; \exists \text{ Pfad } x = (x_0, \ldots x_T) \text{ mit } x_0 = \xi \text{ und } x_T = \eta.$$

Ein Pfad heißt Zirkel (loop)*, falls $x_0 = x_T$ gilt. Hat $(\mathfrak{X}, \prec)$ Zirkel, so ist $\overset{\mathrm{T}}{\prec}$ nicht asymmetrisch.*

Definition 1.1.3. *Ein* Baum *ist ein Graph $(\mathfrak{X}, \prec)$ mit:*

1) $\prec$ und $\overset{\mathrm{T}}{\prec}$ sind asymmetrisch (also gibt es keine loops)

2) $\exists$ ausgezeichnetes $\xi_0 \in \mathfrak{X}$ mit

$$\xi_0 \overset{\mathrm{T}}{\prec} \eta \quad \forall\, \eta \in \mathfrak{X}, \eta \neq \xi_0$$

3) $\forall\, \eta \neq \xi_0$ gilt $|\{\xi \mid \xi \prec \eta\}| = 1$

Wir schreiben $V(\eta) := \{\xi \mid \xi \prec \eta\}$ für die Menge der Vorgänger von $\eta \in \mathfrak{X}$. Ähnlich $N(\eta) := \{\xi \mid \eta \prec \xi\}$ für die Nachfolger. Ferner sei

$$\partial\mathfrak{X} := \{\eta \mid N(\eta) = \emptyset\}$$

der *Rand* von $\mathfrak{X}$. Für ξ_0 ist $V(\xi_0) = \emptyset$, sonst $|V(\xi)| = 1$.

Satz 1.1.4. *Zu jedem $\xi \in \mathfrak{X}$ gibt es genau einen Pfad, der ξ_0 und ξ verbindet.*

Beweis. Klar. □

Definition 1.1.5. *Sei $\xi_0^* \in \mathfrak{X}$ und*

$$\mathfrak{X}^* := \{\xi \in \mathfrak{X} \mid \xi_0^* \overset{\mathrm{T}}{\prec} \xi\} \cup \{\xi_0^*\}$$

$$\prec^* := \prec \cap (\mathfrak{X} \times \mathfrak{X})$$

Dann heißt $(\mathfrak{X}^*, \prec^*) = (\mathfrak{X}^{\xi_0^*}, \prec^{\xi_0^*})$ der von ξ_0^* generierte Teilbaum.
(Man zeigt leicht: $(\mathfrak{X}^*, \prec^*)$ ist Baum.)

Hier ist das ENDE der Beschreibung des Dynamischen Systems.
Beschreibe nun „Incentives" für Kontrolle.

Definition 1.1.6. *Sei $(\mathfrak{X}, \prec)$ ein Baum. Eine Funktion*

$$f\colon \prec \to \mathbb{R}$$

heißt (intermediäre, temporäre, Zwischen-) Auszahlung *(*reward*). Eine Funktion*

$$u\colon \partial\mathfrak{X} \to \mathbb{R}$$

heißt (finale, End-) Auszahlung.

Definition 1.1.7. $\Sigma = (\mathfrak{X}, \prec, f, u)$ *ist ein dynamisches* Entscheidungsproblem *(Entscheidungsbaum).*

ENDE der Beschreibung der Incentives. Nun: Was ist strategisches Verhalten?
Wähle Pfad, nach Möglichkeit optimal. (nur zwischenzeitlich – „Strategie" soll später anders definiert werden).
Ein Pfad (x_0, x_T) mit $x_0 = \xi_0,\ x_T \in \partial\mathfrak{X}$ nennen wir eine *Partie*. Sei

$$\underline{\overline{\mathsf{X}}} = \{x = (x_0, \dots, x_T) \mid x \text{ ist Partie}\}$$

Die Funktion

$$\begin{aligned} &C^0\colon \underline{\overline{\mathsf{X}}} \to \mathbb{R} \\ &C^0(x) = f(x_0, x_1) + f(x_1, x_2) + \cdots + f(x_{T-1}, x_T) + u(x_T) \end{aligned}$$

ist die *Evaluierung* (der Partien).

Definition 1.1.8. *Sei Σ ein dynamisches Entscheidungsproblem (ein E-Baum). Das Paar*

$$\Gamma^0 = \Gamma^0_\Sigma := (\underline{\overline{\mathsf{X}}}, C^0)$$

ist die abgeleitete Normalform. *$x \in \underline{\overline{\mathsf{X}}}$ heißen Alternativen, C^0 die Evaluierung. Der* Wert *von Γ^0 ist*

$$v_0(\xi_0) := \max\{C^0(x) \mid x \in \underline{\overline{\mathsf{X}}}\}$$

Eine Partie $\bar{x}$ ist optimal, *wenn $C^0(\bar{x}) = v_0(\xi_0)$ gilt.*

Wir diskutieren nun das

1.1.1 Optimalitätsprinzip

welches besagt:

> Jeder Teilpfad eines optimalen Pfades ist optimal.
> (Bellman, Isaacs, Zermelo-v. Neumann-Kuhn, ..., Euklid)

Genauer: Zunächst wird die Evaluierung fortgesetzt auf beliebige Pfade: Für jeden Pfad $x = (\xi, x_1, \dots, x_S)$, der $\xi \in \mathfrak{X}$ und $\partial\mathfrak{X}$ verbindet, setzt man

$$C^{0\xi}(x) := f(\xi, x_1) + \cdots + f(x_{S-1}, x_S) + u(s_S)$$

und der *Wert bei Start in* ξ ist

$$v_0(\xi) := \max\{C^{0\xi}(x) \mid x \text{ verbindet } \xi \text{ und den Rand}\}$$

Ein optimaler Pfad bei Start in ξ ist x mit $v_0(\xi) = C^{0\xi}(x)$.

Satz 1.1.9 (Optimalitätsprinzip). *Sei Σ ein dynamischer E-Pfad. Dann ist*

$$x = (\xi, x_1, \dots, x_S)$$

optimal bei ξ, genau dann wenn für jedes x_t von x auch $(x_t, \dots, x_S)$ optimal ist.

Beweis. Sei t und x_t aus einem optimalen Pfad x fixiert. Angenommen, $\tilde{x} = (x_t, \dots, x_S)$ ist nicht optimal. Dann $\exists \hat{x} = (x_t, \hat{x}_{t+1}, \dots, \hat{x}_R)$, der besser ist, d. h. $C^{0x_t}(\hat{x}) > C^{0x_t}(\tilde{x})$. Dann ist aber

$$\begin{aligned}
C^{0\xi}(x) &= f(\xi, x_1) + \cdots + f(x_{S-1}, x_S) + u(x_S) \\
&= f(\xi, x_1) + \cdots + C^{0x_t}(\tilde{x}) \\
&< f(\xi, x_1) + \cdots + C^{0x_t}(\hat{x}) \\
&= f(\xi, x_1) + \cdots + f(x_t, \hat{x}_{t+1} + \cdots + f(\hat{x}_{R-1}, \hat{x}_R) + u(\hat{x}_R) \\
&=: C^{0\xi}(x^*)
\end{aligned}$$

mit einem offensichtlichen x^*, das ξ und $\partial\mathfrak{X}$ verbindet! $\lightning$

$\square$

Bemerkung 1.1.10. Wir haben gesehen: für jedes $\xi^* \in \mathfrak{X}$ ist ein Teilbaum

$$(\mathfrak{X}^*, \prec^*) = (\mathfrak{X}^{\xi^*}, \prec^{\xi^*})$$

definiert. Mit $f^* := f|_{\prec^*}$ und $u^* := u|_{\partial\mathfrak{X}^*}$ ist auch

$$\Sigma^* := \Sigma^{\xi^*} = (\mathfrak{X}^*, \prec^*, f^*, u^*)$$

definiert. Die Wertfunktion

$$v_0 \colon \mathfrak{X} \to \mathbb{R}$$

hat folgende Bedeutungen:

- $v_0(\xi)$ ist maximale Auszahlung eines Pfades in Σ, der bei ξ startet.

- $v_0(\xi)$ ist maximale Auszahlung eines Pfades in Σ, der bei ξ_0 startet, gerechnet ab ξ.
- $v_0(\xi)$ ist maximale Auszahlung des Dynamischen E-Problems Σ^ξ, also $v_{\Sigma^\xi}(\xi)$

Man hat

$$v_{0\Sigma^\xi}(\eta) = v_{0\Sigma^{\xi_0}}(\eta) = v_0(\eta)$$

für $\xi_0 \overset{\mathrm{T}}{\prec} \xi \prec \eta$. Man hat eine Familie von Wertfunktionen!!!

Satz 1.1.11 (Die Optimalitätsgleichung). *Sei Σ ein dynamisches E-Problem. Die Wertfunktion $v\colon \mathfrak{X} \to \mathbb{R}$ erfüllt:*

1) $v_0(\xi) = \max_{\eta \in N(\xi)}\{f(\xi,\eta) + v_0(\eta)\} \quad \forall\, \xi \in \mathfrak{X} \setminus \partial\mathfrak{X}$

2) $v_0(\xi) = u(\xi) \quad \forall\, \xi \in \partial\mathfrak{X}$

1) heißt Optimalitätsgleichung, 2) ist die Randbedingung. 1) bedeutet: Man beachte die Kosten eines Schrittes und den Wert „von da an“ und maximiere die Summe beider.

Beweis. 2) ist klar. Wir zeigen also 2). Sei dazu $\bar{x} = (\xi, \bar{x}_1, \dots, \bar{x}_T)$ ein optimaler Pfad bei ξ. Nach dem Optimalitätsprinzip (Satz 1.1.9) ist $(\bar{x}_2, \dots, \bar{x}_T)$ optimal für Start bei $\bar{x}_1$. Mithin folgt

$$\begin{aligned} v_0(\xi) &= f(\xi, \bar{x}_1) + f(\bar{x}_1, \bar{x}_2) + \dots + f(\bar{x}_{T-1}, \bar{x}_T) + u(\bar{x}_T) \\ &= f(\xi, \bar{x}_1) + v_0(\bar{x}_1) \end{aligned}$$

Sei ferner $(\xi, x_1, \dots, x_S)$ irgendein Pfad von ξ zum Rand, dann ist

$$v_0(\xi) \geq f(\xi, x_1) + f(x_1, x_2) + \dots + u(x_S)$$

Mithin ist $\forall\, \eta \in N(\xi)$ stets

$$\begin{aligned} v_0(\xi) &\geq f(\xi,\eta) + \max\{f(\eta, x_2) + \dots + u(x_S) \mid (\eta, x_2, \dots, x_S) \text{ verbindet } \eta \text{ und } \mathfrak{X}\} \\ &= f(\xi,\eta) + v_0(\eta), \end{aligned}$$

also folgt Gleichheit. □

Bemerkung 1.1.12. v_0 ist eindeutig durch die beiden Bedingungen 1) und 2). Denn v_0 ist auf $\partial\mathfrak{X}$ eindeutig (da $= u$). Ist v_0 aber eindeutig auf allen Nachfolgern von $\xi \in \mathfrak{X} - \partial X$, so auch auf ξ wegen 1).
Betrachtet man einen optimalen Pfad, so sieht man

$$\begin{aligned} v_0 &= \max_{\eta \in N(\xi)} \{f(\xi,\eta) + v_0(\eta)\} \\ &= f(\xi, \bar{x}_1) + v_0(\bar{x}_1) \end{aligned}$$

D. h., der Anfang eines optimalen Pfades ist ein Maximierer der Optimalitätsgleichung! Umgekehrt auch: $\hat{\eta} \in \mathsf{M}_{N(\xi)}\{f(\xi,\bullet) + v_0(\bullet)\}$ ist Beginn eines optimalen Pfades! Ist nämlich $(\xi, \eta, x_2, \dots, x_R)$ ein beliebiger Pfad von ξ nach $\partial\mathfrak{X}$, so folgt für $\hat{\eta} \in \mathsf{M}$:

$$\begin{aligned} f(\xi,\eta) + f(\eta, x_2) + \dots + u(x_R) &\leq f(\xi,\eta) + v_0(\eta) \\ &\leq \max_{\eta'}\{f(\xi,\eta') + v_0(\eta')\} \\ &= f(\xi,\hat{\eta}) + v_0(\hat{\eta}) \\ &= f(\xi,\hat{\eta}) + f(\hat{\eta}, \hat{x}_2) + \dots + u(\hat{x}_S) \end{aligned}$$

d. h., $\hat{\eta}$ beginnt den Pfad. Also: Die Maximierer der Optimalitätsgleichung sind genau die Anfangsknoten optimaler Pfade!

Bemerkung 1.1.13. Zusammenfassend kann man folgende Prozedur beschreiben, die Wertfunktion und optimale Pfade liefert. ("Backwards Induction")

1) Wegen 2) ist $v_0(\xi) = u(\xi)$ für $\xi \in \partial\mathfrak{X}$

2) Hat man v_0 auf allen $\eta \in N(\xi)$ bestimmt, so berechnet man $v_0(\xi)$ vermöge 1). Man markiert gleichzeitig einen Maximierer (also optimale Richtung, Beginn eines optimalen Pfades).

3) Hat man für jedes $\eta \in N(\xi)$ einen optimalen Pfad, so liefert der Maximierer $\bar{\eta}$ zusammen mit dem optimalen Pfad einen optimalen Pfad von ξ.

4) Graphisch: markiere optimale Nachfolger sukzessive durch Pfeile. Notiere $v_0(\xi)$ ebenfalls sukzessive. Laufe dann von der Wurzel entlang der Pfeilrichtungen – dies liefert optimalen Pfad.

Beispiel 1.1.14. Zur Darstellung des Algorithmus folgendes Bild: Abb. (1.1) zeigt ein Op-

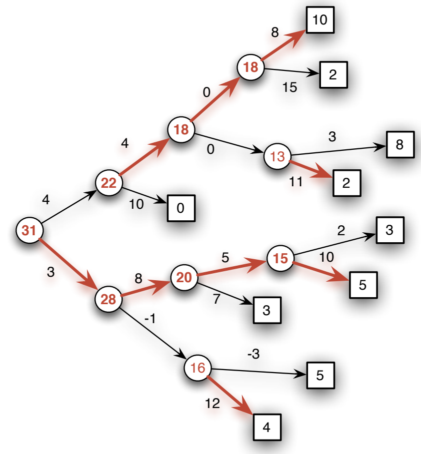


Abbildung 1.1: Darstellung der Rückwärtsinduktion ("Backwards Induction")

timierungsproblem mit $v_0 = 31$.

Wir ändern unter dem Eindruck der *rekursiven Methode* unseren Begriff von *strategischem Verhalten* (bisher: Pfade)

Definition 1.1.15. *Sei $\Sigma = (\mathfrak{X}, \prec, f, u)$ ein Entscheidungsbaum (dynamisches Optimierungsproblem). Eine* Strategie *(oder* Kontrolle*) ist eine Abbildung*

$$\alpha \colon \mathfrak{X} \setminus \partial\mathfrak{X} \to \mathfrak{X}$$

derart, daß

$$\alpha(\xi) \in N(\xi)$$

für alle $\xi \in \mathfrak{X} \setminus \partial\mathfrak{X}$ gilt.

Jede Strategie erzeugt einen Pfad $X(\alpha) = X^{\xi_0}(\alpha)$ vermöge der Festsetzung $X(\alpha) = (x_0, x_1, \ldots, x_T)$ mit $x_0 = \xi_0$, $x_1 = \alpha(x_0)$,...,$x_T = \alpha(x_{T-1})$, so daß $x_0 \prec x_1 \prec \ldots \prec x_T$ erfüllt ist. Zusammenfassend:

Definition 1.1.16. *Sei Σ ein dyn. Opt.-Problem.*

1) $\mathfrak{S} = \{\alpha \mid \alpha\colon \mathfrak{X} \setminus \partial\mathfrak{X} \to \mathfrak{X},\ \alpha(\xi) \in N(\xi)\ \forall \xi \in \mathfrak{X} \setminus \partial\mathfrak{X}\}$ *ist die Menge aller reinen Strategien.*

2) *Die Funktion* $C_\bullet = C_\bullet^{\xi_0}\colon \mathfrak{S} \to \mathbb{R}$, $C_\alpha^{\xi_0} = C^{0\xi}(X(\alpha))$ *ist die* Auszahlungsfunktion (payoff function).

3) $v(\xi_0) = \max\{C_\alpha^{\xi_0} \mid \alpha \in \mathfrak{S}\}$ *ist der* Wert *des Entscheidungsproblems* $(\mathfrak{S}, C_\bullet^{\xi_0})$.

Bemerkung 1.1.17. Natürlich gibt es zu jedem $\xi \in \mathfrak{X}$ ein Σ^ξ und eine Normalform

$$\Gamma_{\Sigma^\xi} = (\mathfrak{S}, C_\bullet^\xi),$$

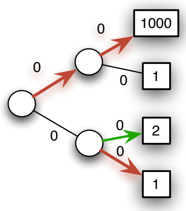
dabei spielt es keine wesentliche Rolle, ob man Strategien $\alpha\colon \mathfrak{X} \setminus \partial\mathfrak{X} \to \mathfrak{X}$ oder ihre Restriktionen $\alpha|_{\mathfrak{X}^* - \partial\mathfrak{X}^*}$ oder Strategien $\alpha\colon \mathfrak{X}^* \setminus \partial\mathfrak{X}^* \to \mathfrak{X}^*$ einsetzt. Unser Prinzip des Übergangs ist stets

1) Definiere die Extensive Form (dyn. Prozeß, E-Baum, Spielregeln,...)

2) Definition strategischen Verhaltens und Auszahlungsfunktion

3) Aufstellung der Normalform

Für Strategien gilt das Optimalitätsprinzip: Eine Strategie α ist optimal genau dann, wenn sie optimal ist für jeden Teilbaum, den man bei Anwendung von α erreicht. (Nicht notwendig bei jedem Teilbaum):



Satz 1.1.18. *1) Sei $\bar\alpha$ eine optimale Strategie für Start bei $\xi \in \mathfrak{X}$ und sei $\bar x = X(\bar\alpha)$ der erzeugte Pfad. Dann ist $\bar\alpha$ optimal bei Start von jedem $\bar x_0, \ldots, \bar x_T$.*

2) Die Funktion

$$\begin{aligned} v&\colon \mathfrak{X} \to \mathbb{R} \\ v(\xi) &= \max\{C_\alpha^\xi \mid \alpha \in \mathfrak{S}\} \end{aligned}$$

leistet

(1) $v(\xi) = \max_{\eta \in N(\xi)}\{f(\xi,\eta) + v(\eta)\}\ \forall \xi \in \mathfrak{X} \setminus \partial\mathfrak{X}$
(2) $v(\xi) = u(\xi)\ \forall \xi \in \partial\mathfrak{X}$

3) Ist für jedes $\xi \in \mathfrak{X} \setminus \partial\mathfrak{X}$ *in (2.1) ein Maximierer* $\bar{\alpha}(\xi)$ *bestimmt, so ist das so erzeugte* $\bar{\alpha}\colon \mathfrak{X} \setminus \partial\mathfrak{X} \to \mathfrak{X}$ *eine optimale Strategie für jeden Teilbaum!*

Bemerkung 1.1.19. $v = v_0$, da (2.1) und (2.2) eindeutige Lösungen haben!

1.2 Spiele über einem Baum

Wir folgen dem Schema wie eben: zugrunde liegt zusätzlich eine *Spielermenge*

$$I = \{1, \dots, n\}$$

Definition 1.2.1. *Ein extensives Spiel in Baum-Form (ein* Spielbaum*) ist ein Tupel*

$$\Sigma = (\mathfrak{X}, \prec, \iota, f, u)$$

mit folgenden Daten:

1) $(\mathfrak{X}, \prec)$ *ist ein Baum*

2) $\iota\colon \mathfrak{X} \setminus \partial\mathfrak{X} \to I$ *ist eine Abbildung,* ι *ist die* Spielerzuweisung*. D. h., kommt der Pfad bei* ξ *an und ist* $\iota(\xi) = i \in I$*, so entscheidet Spieler* i *über die Auswahl des Nachfolgers. Wir nehmen stets an, daß* ι *surjektiv ist (jedes* i *darf mindestens einmal entscheiden).*

3) $f = (f^i)_{i \in I}$, $f^i\colon \prec \to \mathbb{R}$. f^i *ist die intermediäre Auszahlung an Spieler* i*.*

4) $u = (u^i)_{i \in I}$, $u^i\colon \partial\mathfrak{X} \to \mathbb{R}$. u^i *ist die terminale Auszahlung an Spieler* i*.*

Beispiel 1.2.2. Wir repräsentieren Σ als Graphen:

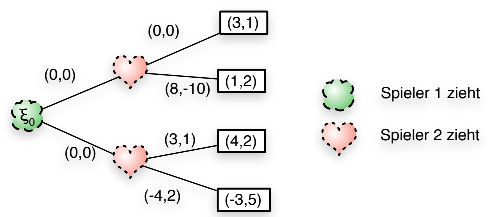


Abbildung 1.2: Baum mit zwei Spielern

Bemerkung 1.2.3. Wir schreiben

$$\begin{aligned} \mathfrak{X}^i &:= \{\xi \in \mathfrak{X} \mid \iota(\xi) = i\} \\ &= \iota^{-1}(\{i\}) \end{aligned}$$

für die von i regierten Knoten, stets ist $\mathfrak{X}^i \neq \emptyset$. Dies erzeugt eine Zerlegung von $\mathfrak{X} \setminus \partial\mathfrak{X}$:

$$\mathfrak{X} \setminus \partial\mathfrak{X} = \dot{\bigcup}_{i \in I} \mathfrak{X}^i$$

Die Evaluierung ist wie früher, jedoch für jeden Spieler definiert:

$$\begin{aligned} C^i(x) = C^{i\xi_0}(x) &= C^{i\xi_0}(x_0, \dots, x_T) \\ &= f^i(x_0, x_1) + \dots + f^i(x_{T-1}, x_T) + u^i(x_T) \end{aligned}$$

für Pfade $x = (x_0, \dots, x_T)$ mit $x_0 = \xi_0$ und $x_{t-1} \prec x_t \ \forall t \in \{1, \dots, T\}$

Definition 1.2.4. *Eine (reine) Strategie für Spieler $i \in I$ ist eine Abbildung*

$$\begin{aligned} &\alpha^i \colon \mathfrak{X}^i \to \mathfrak{X} \\ \textit{mit} \qquad &\alpha^i(\xi) \in N(\xi) \qquad \forall \xi \in \mathfrak{X}^i \end{aligned}$$

Anschaulich: ein Pfeil bei allen Knoten, die i regiert.

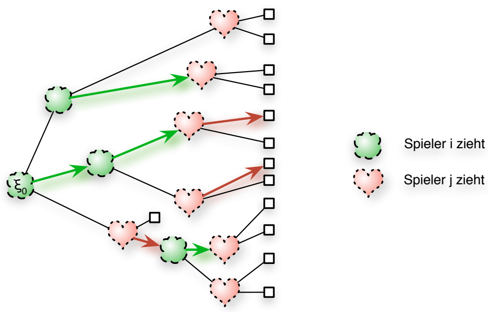


Abbildung 1.3: Baum mit zwei Spielern und Strategien

Natürlich induziert jedes Strategie-n-Tupel

$$\alpha = (\alpha^1, \dots, \alpha^n) = (\alpha^i)_{i \in I}$$

einen Pfad

$$X(\alpha) = (x_0, x_1, \dots, x_T)$$

wie früher:

$$\begin{aligned} x_0 &= \xi_0 \\ x_1 &= \alpha^i(\xi_0) \text{ falls } \xi_0 \in \mathfrak{X}^i \\ &\vdots \\ x_T &= \alpha^j(x_{T-1}) \text{ falls } x_{T-1} \in \mathfrak{X}^j \end{aligned}$$

Damit ist die Auszahlung auf Strategie-n-Tupel definiert:

$$C^{i\xi}_{\alpha} = C^{i\xi}_{(\alpha^i, \dots, \alpha^n)} = C^{i\xi}(X(\alpha))$$

für Strategie-n-Tupel α.

Definition 1.2.5 (Die Normalform). *Sei* $\mathfrak{S}^i := \{\alpha^i \mid \alpha^i$ *ist Strategie von* $i\}$ *die von* $\Sigma = \Sigma^\xi$ *abgeleitete* Normalform *ist das (nichtkooperative)* n*-Personen-Spiel*

$$\Gamma = \Gamma_{\Sigma^\xi} = \Gamma^\xi_\Sigma = (\mathfrak{S}^1, \ldots, \mathfrak{S}^n, C^{1\xi}_\bullet, \ldots, C^{n\xi}_\bullet)$$

mit $C^{i\xi}_\bullet : \mathfrak{S}^1 \times \cdots \times \mathfrak{S}^n \to \mathbb{R}\ \forall i \in I$.

Bemerkung 1.2.6.

- Wir unterscheiden zwischen *Aktion* (lokale Entscheidung an einem Knoten) und *Strategie* (globaler Plan).
- Für $n = 2$ ist Γ_{Σ^ξ} ein Bimatrixspiel in reinen Strategien:

$$\Gamma = (\mathfrak{S}^1, \mathfrak{S}^2, C^{1\xi}_\bullet, C^{2\xi}_\bullet) \qquad \mathfrak{S}^1, \mathfrak{S}^2 \text{ sind endlich}$$

- Für jedes $\Sigma = \Sigma^{\xi_0}$ gibt es eine ganze Familie $\Sigma^\xi\ \forall \xi \in \mathfrak{X}$ von Spielbäumen und dementsprechend eine Familie von Normalformen

$$\Gamma_{\Sigma^\xi} = \Gamma^\xi_\Sigma$$

Beispiel 1.2.7. t, r, l etc suggerieren Aktionen. Man hat:

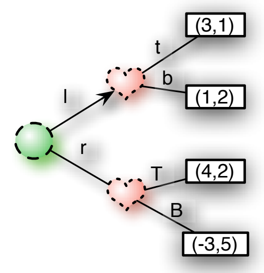


Abbildung 1.4: Baum als Bimatrixspiel

$$\mathfrak{S}^1 = \{l, r\} \qquad\qquad \mathfrak{S}^2 = \{tT, tB, bT, bB\}$$

und daher

$$C^{1\xi_0}_\bullet = \begin{pmatrix} 3 & 3 & 1 & 1 \\ 4 & -3 & 4 & -3 \end{pmatrix} \qquad C^{2\xi_0}_\bullet = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 2 & 5 & 2 & 5 \end{pmatrix}$$

Man bewertet ein Nash-Gleichgewicht bei (l, bB).

Bemerkung 1.2.8.

1) Wie üblich hat man eine ganze Familie von Spielen $\Sigma^\xi, \Gamma^\xi_\Sigma$ für $\xi \in \mathfrak{X}$. (Stelle die Normalform für Σ^l im Bsp. auf!)
2) Jede Strategie kann auch in Σ^ξ benützt werden.
3) Ist α ein Strategie-n-Tupel, so sind die Auszahlungen in der Gegenwart und in der Zukunft verknüpft durch

$$C^{l\xi}_\alpha = C^{l,\alpha^i(\xi)}_\alpha + f^l(\xi, \alpha^i(\xi))$$

für $i, l \in I$ und $\xi \in \mathfrak{X}^i$.

Definition 1.2.9 (Backwards Induction). *Sei Σ ein Spielbaum und $\bar{\alpha} = (\bar{\alpha}^1, \dots, \bar{\alpha}^n)$ ein Strategie-n-Tupel. Man sagt, $\bar{\alpha}$ ist durch Rückwärtsinduktion gewonnen („$\bar{\alpha}$ ist teilspielperfekt“), wenn es eine Familie $v^i : \mathfrak{X} \to \mathbb{R}$ gibt mit folgenden Eigenschaften:*

1) Für $i, l \in I$ und $\xi \in \mathfrak{X}^i$:

$$\begin{aligned} v^i(\xi) &= \max_{\eta \in N(\xi)} \{v^i(\eta) + f^i(\xi, \eta)\} \\ &= v^i(\bar{\alpha}^i(\xi)) + f^i(\xi, \bar{\alpha}^i(\xi)) && (1.1) \\ v^l(\xi) &= v^l(\bar{\alpha}^i(\xi)) + f^l(\xi, \bar{\alpha}^i(\xi)) && (1.2) \\ v^i(\xi) &= u^i(\xi) \qquad \xi \in \partial\mathfrak{X} && (1.3) \end{aligned}$$

(1.1) und (1.2) sind Opt.gleichungen, (1.3) ist Randbedingung.

Satz 1.2.10. *Sei Σ ein Spielbaum und Γ_Σ^ξ die Familie der Normalformen. Sei $\bar{\alpha}$ durch Rückwärtsinduktion gewonnen. Dann ist für jedes ξ stets $\bar{\alpha}$ (genauer: die Restriktion) Nash-Gleichgewicht in Γ_Σ^ξ („Teilspielperfekt“). Ferner ist $v^i(\xi)$ die Gleichgewichts-Auszahlung an i bei Start in ξ.*

Beweis. Durch Induktion nach der maximalen Pfadlänge, für $\mathfrak{X} = \{\xi_0\}$ ist nichts zu zeigen. Sei $i \in I$ und $\xi \in \mathfrak{X}^i$. Für alle $l \in I$ gilt dann:

$$\begin{aligned} v^l(\xi) &= v^l(\bar{\alpha}^i(\xi)) + f^l(\xi, \bar{\alpha}^i(\xi)) && (1.1) \\ &= C^{l, \bar{\alpha}^i(\xi)} + f^l(\xi, \bar{\alpha}^i(\xi)) && \text{Induktionsvoraussetzung} \\ &= C_{\bar{\alpha}}^{l\xi} && \text{nach Bemerkung} \end{aligned}$$

In der gleichen Situation (i regiert ξ) gilt: Weicht i von $\bar{\alpha}^i$ ab zu $\hat{\alpha}^i$, so folgt

$$\begin{aligned} C^{i\xi} &= v^i(\xi) \\ \bar{\alpha} &= v^i(\bar{\alpha}^i(\xi)) + f^i(\xi, \bar{\alpha}^i(\xi)) \\ &\geq v^i(\hat{\alpha}^i(\xi)) + f^i(\xi, \hat{\alpha}^i(\xi)) \\ &= C_{\bar{\alpha}}^{i, \hat{\alpha}^i(\xi)} + f^i(\xi, \hat{\alpha}^i(\xi)) \\ &\geq C_{(\hat{\alpha}^i, \bar{\alpha}^{-i})}^{i, \hat{\alpha}^i(\xi)} + f^i(\xi, \hat{\alpha}^i(\xi)) && \text{nach Induktion} \\ &= C_{(\hat{\alpha}^i, \bar{\alpha}^{-i})}^{i\xi} && \text{nach Bemerkung} \end{aligned}$$

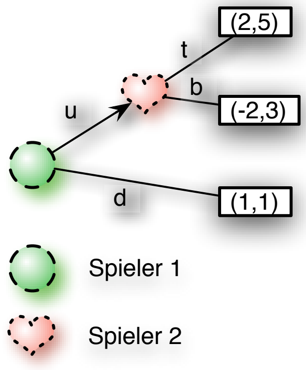
Also ist $\bar{\alpha}$ auch Gleichgewicht bei Start in ξ und – induktiv – überall. □

Korollar 1.2.11. Jedes Baumspiel hat (teilspielperfekte) Nash-Gleichgewichte.

Beispiel 1.2.12.

$$C_\bullet^{1\xi_0} = \begin{pmatrix} 2 & -2 \\ 1 & 0 \end{pmatrix} \qquad C_\bullet^{2\xi_0} = \begin{pmatrix} 5 & 3 \\ 1 & 1 \end{pmatrix}$$

(u, t) ist teilspielperfektes Gleichgewicht. Auch (d, b) ist ein (nicht teilspielperfektes) Gleichgewicht! Spieler 2 stabilisiert dieses Gleichgewicht, weil Spieler 2 „androht“, b zu wählen, falls Spieler 1 u wählen sollte. Die Drohung ist aber nicht glaubwürdig: sollte Spieler 1 dennoch u wählen, ist es für Spieler 2 nicht rational, b zu wählen, da t ihm höhere Auszahlung bringt!

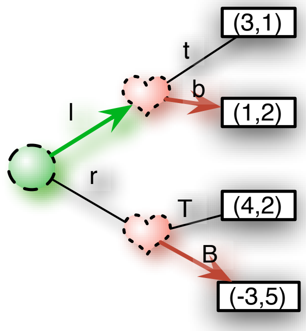


Nicht-teilspielperfekte Gleichgewichte beinhalten unglaubwürdige Drohungen.

Beispiel 1.2.13.

$$C_\bullet^{1\xi_0} = \begin{pmatrix} 3 & 3 & 1 & 1 \\ 4 & -3 & 4 & -3 \end{pmatrix} \qquad C_\bullet^{2\xi_0} = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 2 & 5 & 2 & 5 \end{pmatrix}$$

das teilspielperfekte Gleichgewicht ist (l, bB) mit Auszahlung



$$(v^1(\xi_0), v^2(\xi_0)) = (1, 2)$$

Bemerkung 1.2.14. Sei $n = 2$ und $u^2 = -u^1$, $f^2 = -f^1$. Dann ist auch

$$C_\bullet^\xi := C_\bullet^{1\xi} = -C_\bullet^{2\xi}$$

(d. h., Γ_Σ^ξ ist stets Nullsummenspiel.) Wir lassen dann stets die Indizes ^{1,2} entfallen. D. h., auch

$$\Gamma = \Gamma_\Sigma = \Gamma_\Sigma^\xi = (\mathfrak{S}^1, \mathfrak{S}^2, C_\bullet^\xi)$$

ist ein Matrix-(Nullsummen)spiel. In diesem Falle ist ein Nash-Gleichgewicht äquivalent zu einem Sattelpunkt, d. h., $\bar{\alpha}$ mit

$$C^\xi_{\bar{\alpha}^1, \alpha^2} \geq C^\xi_{\bar{\alpha}^1, \bar{\alpha}^2} \geq C^\xi_{\alpha^1, \bar{\alpha}^2} \qquad \forall (\alpha^1, \alpha^2) \in \mathfrak{S}^1 \times \mathfrak{S}^2$$

Nullsummenspiele haben (falls es Sattelpunkte=Nash-Gleichgewichte gibt) Wert und optimale Strategie, d. h.

$$\begin{aligned} v_{\Gamma^\xi_\Sigma} &= \operatorname*{val}_{\mathfrak{S}^1\times\mathfrak{S}^2} C^\xi_\bullet \\ &= \max_{\alpha^1\in\mathfrak{S}^1}\min_{\alpha^2\in\mathfrak{S}^2} C^\xi_{\alpha^1,\alpha^2} \\ &= \min_{\alpha^2\in\mathfrak{S}^2}\max_{\alpha^1\in\mathfrak{S}^1} C^\xi_{\alpha^1,\alpha^2} \\ &= C^\xi_{\bar\alpha} =: v(\xi) \end{aligned}$$

Korollar 1.2.15. Ist Σ und Γ^ξ_Σ ein Nullsummenspiel, dann besitzt Γ^ξ_Σ für jedes ξ Wert und optimale Strategien. Die (Spiel-) Wertfunktion erfüllt die Gleichungen der Rückwärtsinduktion

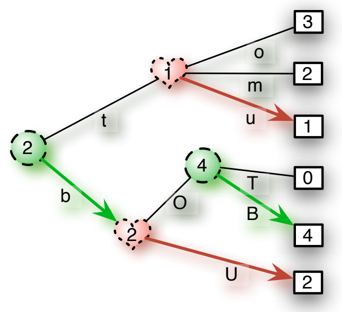
$$\begin{aligned} v(\xi) &= \max_{\eta\in N(\xi)}\{v(\eta)+f(\xi,\eta)\} \\ &= v(\bar\alpha^1(\xi))+f(\xi,\bar\alpha^1(\xi)) \qquad \forall\xi\in\mathfrak{X}^1 \\ v(\xi) &= \min_{\eta\in N(\xi)}\{v(\eta)+f(\xi,\eta)\} \qquad \forall\xi\in\mathfrak{X}^2 \\ &= v(\bar\alpha^2(\xi))+f(\xi,\bar\alpha^2(\xi)) \end{aligned}$$

sowie $v(\xi) = u(\xi)\ \forall\xi \in \partial\mathfrak{X}$ und diese definieren optimale Strategien $\bar\alpha^1, \bar\alpha^2$ für beide Spieler.

Beispiel 1.2.16.

$$C^\xi_\bullet = \begin{array}{c|cccccc} & oO & oU & mO & mU & uO & uU \\ \hline tT & 3 & 3 & 2 & 2 & 1 & 1 \\ tB & 3 & 3 & 2 & 2 & 1 & 1 \\ bT & 0 & 2 & 0 & 2 & 0 & 2 \\ bB & 4 & 2 & 4 & 2 & 4 & 2 \end{array}$$

optimal ist hier: (bB, uU). Also: $v(\Gamma^\xi_\Sigma) = 2$



1.3 Dynamische Spiele, der Operator der Rückwärtsinduktion

Wir legen wieder zugrunde $I = \{1,\ldots,n\}$, die Spielermenge. Wie bisher: 3 Schritte:

1. Schritt:

Extensive Form aufstellen: Es sei $T > 0$, $T \in \mathbb{N}$; T bezeichnet Dauer des Prozesses. Wir schreiben $\mathsf{T}_0 := \{0, \ldots, T\}$, $\mathsf{T}_1 := \{1, \ldots, T\}$. Für jedes $t \in \mathsf{T}_0$ sei $\underline{\mathsf{X}}_t$ eine endliche Menge (die Zustandsmenge des Prozesses zur Zeit t), bekannt bei allen Spielern. Für jedes $i \in I$ und $t \in \mathsf{T}$ sei $\overline{\underline{\mathsf{Y}}}^i_t$ eine endliche Menge (die Aktionen von i zur Zeit t). Für jedes $t \in \mathsf{T}$ sei $a_t \colon \overline{\underline{\mathsf{X}}}_{t-1} \times \underline{\mathsf{Y}}_t \to \underline{\mathsf{X}}_t$ die Übergangsfunktion. Wir setzen $\overline{\underline{\mathsf{Y}}}_t := \overline{\underline{\mathsf{Y}}}^1_t \times \cdots \times \overline{\underline{\mathsf{Y}}}^n_t$ (D. h. für $\xi \in \overline{\underline{\mathsf{X}}}_{t-1}$ und $\eta = (\eta^1, \ldots, \eta^n) \in \underline{\mathsf{Y}}_t$ ist $a_t(\xi, \eta) \in \underline{\mathsf{X}}_t$ der nächste Zustand.)
Dei Rewards: Für jedes $i \in I$ sei $f^i_t \colon \underline{\mathsf{X}}_{t-1} \times \underline{\mathsf{Y}}_t \to \mathbb{R}\ \forall t \in \mathsf{T}$ der intermediäre Nutzen für Spieler i zur Zeit t. Zu jedem Zustand $\xi \in \overline{\underline{\mathsf{X}}}_{t-1}$ ist eine n-dimensinale Matrix $f^\bullet_t(\xi, \bullet)$ gegeben. Ferner sei für $i \in I$, $u^i \colon \overline{\underline{\mathsf{X}}}_T \to \mathbb{R}$ gegeben. Also

Definition 1.3.1. *Ein dynamsiches Spiel (in extensiver Form) ist ein Tupel*

$$\Sigma = (\mathfrak{X}, \mathfrak{Y}, a, f, u, \mathsf{T})$$

mit $\mathfrak{X} = \left(\overline{\underline{\mathsf{X}}}^i_t\right)^{i \in I}_{t \in \mathsf{T}_0}$. *Die Evaluierung von Pfaden ist kanonisch: mit* $\overline{\underline{\mathsf{X}}} := \overline{\underline{\mathsf{X}}}_0 \times \cdots \times \overline{\underline{\mathsf{X}}}_T$ *und* $\underline{\mathsf{Y}} := \overline{\underline{\mathsf{Y}}}_1 \times \cdots \times \overline{\underline{\mathsf{Y}}}_T$ *ist*

$$C^i \colon \overline{\underline{\mathsf{X}}} \times \overline{\underline{\mathsf{Y}}} \to \mathbb{R}$$

$$C^i(x, y) = \sum_{t=1}^{T} f^i_t(x_{t-1}, y_t) + u^i(x_T)$$

2. Schritt

Strategisches Verhalten: Eine reine Strategie für Spieler i ist eine Familie von Abbildungen $\alpha^i = (\alpha^i_1, \ldots, \alpha^i_T)$

$$\alpha^i_t \colon \overline{\underline{\mathsf{X}}}_{t-1} \to \overline{\underline{\mathsf{Y}}}^i_t$$

$\mathfrak{S}^i$ ist die Menge der reinen Strategien, $\mathfrak{S} = \mathfrak{S}^1 \times \cdots \times \mathfrak{S}^n$ die Menge der Strategien-n-Tupel (α^i ist ein globaler Plan für Spieler i).

Definition 1.3.2. *Wir definieren die Abbildung*

$$\begin{array}{ll}
(x^\bullet_*, y^\bullet_*) \colon \overline{\underline{\mathsf{X}}}_0 \times \mathfrak{S} \to \overline{\underline{\mathsf{X}}} \times \overline{\underline{\mathsf{Y}}} & \textit{vermöge} \\
\qquad x^\xi_\alpha = (x_0, \ldots, x_T) & y^\xi_\alpha = (y_1, \ldots, y_T) \\
\qquad x_0 = \xi & y_1 = \alpha_1(x_0) = (\alpha^1_1(x_0), \ldots, \alpha^n_1(x_0)) \\
\qquad x_1 = a(x_0, y_1) & \\
\qquad \vdots & \\
\qquad x_T = a_T(x_{T-1}, y_T) & y_T = \alpha_T(x_{T-1})
\end{array}$$

$(x^\bullet_*, y^\bullet_*)$ *ist der Prozess.*

Die Komposition von Evaluierung und Prozeß definiert nun Auszahlungen über Strategien

n-Tupel:

$$\begin{aligned} C_\alpha^{i\xi} &= u^i(x_T) + \sum_{t=1}^{T} f_t^i(x_{t-1}, \alpha_t(x_{t-1})) \\ &= u^i(x_T) + \sum_{t=1}^{T} f_t^i(x_{t-1}, y_t) \\ &= C^i(x_\alpha^\xi, y_\alpha^\xi) && \text{mit } x_{t-1} = x_{\alpha_{t-1}}^\xi \end{aligned}$$

Daher haben wir nun den

3. Schritt:
Normalform:

Definition 1.3.3. *Die zu Σ (bei ξ) gehörige Normalform ist das n-Personenspiel*

$$\Gamma_\Sigma^\xi = (\mathfrak{S}^1, \dots, \mathfrak{S}^n, C_\bullet^{1\xi}, \dots, C_\bullet^{n\xi})$$

Zur Notation: ^{i} oben, $_t$ unten,

$$\xi^{-i} = (\xi^1, \dots, \xi^{i-1}, \xi^{i+1}, \dots, \xi^n)$$

so daß $\xi = (\xi^i, \xi^{-i})$ so daß z. B. Nash Gleichgewicht in Γ_Σ^ξ ein n-Tupel $\bar\alpha = (\bar\alpha^1, \dots, \bar\alpha^n)$ ist, derart, daß

$$C_{\bar\alpha}^{i\xi} \geq C_{\alpha^i, \bar\alpha^{-i}}^{i\xi} \quad \forall i \in I, \alpha^i \in \mathfrak{S}^i$$

gilt. Zum Rückwärtsrechnen eine weitere Schreibweise: $-s$ statt $T - s$ („komplementär zum Horizont“).
In unserem Kontext haben wir zunächst die Übergangsfunktion

$$a_s \colon \underline{\overline{\mathsf{X}}}_{s-1} \times \underline{\overline{\mathsf{Y}}}_s \to \underline{\overline{\mathsf{X}}}_s$$

Setzt man ein Strategien n-Tupel α ein, so hat man

$$\begin{aligned} a_s^\alpha \colon \underline{\overline{\mathsf{X}}}_{s-1} &\to \underline{\overline{\mathsf{X}}}_x \\ a_s^\alpha(\xi) &= a_s(\xi, \alpha_s(\xi)) && \xi \in \underline{\overline{\mathsf{X}}}_{s-1} \end{aligned}$$

Weicht Spieler i von $\alpha_s^i(\xi)$ zu einem η^i ab, so hat man

$$\begin{aligned} a_s^{\eta^i\alpha} \colon \underline{\overline{\mathsf{X}}}_{s-1} &\to \underline{\overline{\mathsf{X}}}_s \\ a_s^{\eta^i\alpha}(\xi) &= a_s(\xi, (\eta^i, \alpha_s^{-i}(\xi))) \end{aligned}$$

Der typische Term beim Rückwärtsrechnen ist von der Form

$$w(a_{-(s-1)}^\alpha(\xi)) + f_{-(s-1)}^i(\xi, \alpha_{-(s-1)}(\xi))$$

Wir definieren deshalb

$$\mathfrak{F}(\underline{\overline{\mathsf{X}}}_t) = \{w \mid w \colon \underline{\overline{\mathsf{X}}}_t \to \mathbb{R}\}$$

sowie

$$\mathfrak{A}_s^{i\alpha} \colon \mathfrak{F}(\underline{\overline{\mathsf{X}}}_{-(s-1)}) \to \mathfrak{F}(\underline{\overline{\mathsf{X}}}_{-s})$$

mit $(\mathfrak{A}_s^{i\alpha} w)(\xi) = w(a^{\alpha}_{-(s-1)}(\xi)) + f^i_{-(s-1)}(\xi, \alpha_{-(s-1)}(\xi))$. Beachte

$$\begin{aligned} a^{\alpha}_{-(s-1)} &\colon \underline{\overline{X}}_{-s} \to \underline{\overline{X}}_{-(s-1)} \\ \mathfrak{A}_s^{i\alpha} &\colon \mathfrak{F}(\underline{\overline{X}}_{-(s-1)}) \to \mathfrak{F}(\underline{\overline{X}}_{-s}) \end{aligned}$$

$\mathfrak{A}_s^{i\alpha}$ ist der Operator der Rückwärtsinduktion. Wir benötigen noch die Variante, bei der Spieler i zu η^i abweicht.

$$\begin{aligned} \mathfrak{A}_s^{i\eta^i\alpha} &\colon \mathfrak{F}(\underline{\overline{X}}_{-(s-1)}) \to \mathfrak{F}(\underline{\overline{X}}_{-s}) \\ (\mathfrak{A}_s^{i\eta^i\alpha} w)(\xi) &= w(a^{\eta^i\alpha}_{-(s-1)}(\xi)) + f^i_{-(s-1)}(\xi, (\eta^i, \alpha^{-i}_{-(s-1)}(\xi))) \end{aligned}$$

Definition 1.3.4. *Sei Σ ein dynamisches n-Personenspiel in ext. Form und $\bar{\alpha}$ ein Strategie-n-Tupel ($\bar{\alpha} \in \mathfrak{S}$). Man sagt, $\bar{\alpha}$ sei durch Rückwärtsinduktion erhalten oder teilspielperfekt, falls es eine Familie von Funktionen*

$$v^i_t \colon \underline{\overline{X}}_{-t} \to \mathbb{R} \ \forall i \in I, t \in \bar{T}_0$$

gibt, die folgendes leistet:

$$\begin{aligned} v^i_t(\xi) &= \max_{\eta^i \in \underline{\overline{Y}}_{-(t-1)}} \{\mathfrak{A}_t^{i\eta^i\bar{\alpha}} v^i_{t-1}(\xi)\} && (1.4) \\ &= (\mathfrak{A}_t^{i\bar{\alpha}} v^i_{t-1})(\xi) \quad \forall i \in I, t \in T, \xi \in \underline{\overline{X}}_{-t} \\ v^i_0(\xi) &= u^i(\xi) \quad \forall i \in I, \xi \in \underline{\overline{X}}_T && (1.5) \end{aligned}$$

Wie bisher: (1.4) (war Optimalitätsbedingung) ist die rekursive Gleichgewichtsbedingung, (1.5) ist Randbedingung.

Man beachte: die Existenz ist nicht gesichert! Man kann nicht ohne weiteres erwarten, daß die in (1.4) verborgene Gleichgewichtsaufgabe stets lösbar ist: 25. Mai 04

Bemerkung 1.3.5. Die Existenz ist nicht ohne weiteres klar. Betrachte nämlich das „one shot recursive game", das wie folgt gegeben ist:

$$\underline{\overline{Y}}^1_{-(t-1)}, \dots, \underline{\overline{Y}}^n_{-(t-1)}$$

sind die Aktionenräume der Spieler

$$\begin{aligned} H^i = H^{i\xi} &\colon \underline{\overline{Y}}_{-(t-1)} \to \mathbb{R} \\ H^i(\eta) &= v^i_{t-1}(a_{-(t-1)}(\xi, \eta)) + f^i_{-(t-1)}(\xi, \eta) \\ &= (\mathfrak{A}_t^{i\eta} v_{t-1})(\xi) \qquad \eta \in \underline{\overline{Y}}_{-(t-1)} \end{aligned}$$

ist eine Funktion auf $\underline{\overline{Y}}_{-(t-1)}$ und daher ist

$$\Sigma^{\xi}_{-(t-1)} = (\underline{\overline{Y}}^1_{-(t-1)}, \dots, \underline{\overline{Y}}^n_{-(t-1)}, H^1, \dots, H^n)$$

ein nichtkooperatives n-Personenspiel in Normalform. (1.4) besagt, daß dieses ein Nash-Gleichgewicht zuläßt, daß $\bar{\alpha}_{-(t-1)}(\xi)$ eines ist, und daß $v^i_t(\xi)$ eine Gleichgewichtsauszahlung ist. D. h., man muß rückwärtsinduktiv Nash-Gleichgewichte finden! (Man sollte mischen. . .)

Bemerkung 1.3.6. ist Σ ein Nullsummenspiel, d. h., $n = 2$,

$$-f_t^2 = f_t^1 =: f_t$$
$$-u^2 = u^1 =: u$$

So ist auch stets

$$C_\alpha^{1\xi} + C_\alpha^{2\xi} = 0$$

für alle ξ und α, d. h., $\Gamma_\Sigma^\xi = (\mathfrak{S}^1, \mathfrak{S}^2, C_\bullet^\xi)$ ist Nullsummenspiel in Normalform. Das „one shot recursive game“

$$\Gamma_*^\xi = (\underline{\overline{\mathsf{Y}}}_*^1, \underline{\overline{\mathsf{Y}}}_*^2, H)$$

für alle $* = t$ und alle ξ ist auch Nullsummenspiel (induktiv!) Der Spielwert in reinen Strategien ist

$$\begin{aligned} v_{\Gamma_*} &= \operatorname*{val}_{\eta \in \underline{\overline{\mathsf{Y}}}_*} H(\eta) \\ &= \max_{\eta^1 \in \underline{\overline{\mathsf{Y}}}_*^1} \min_{\eta^2 \in \underline{\overline{\mathsf{Y}}}_*^2} H(\eta^1, \eta^2) \\ &= \min_{\eta^2 \in \underline{\overline{\mathsf{Y}}}_*^2} \max_{\eta^1 \in \underline{\overline{\mathsf{Y}}}_*^1} H(\eta^1, \eta^2) \end{aligned}$$

vorausgesetzt, die Größe existiert.

Satz 1.3.7. *Sei Σ ein dyn. Nullsummenspiel. Ein Strategienpaar $\bar{\alpha}$ ist teilspielperfekt genau dann, wenn es eine Familie von Funktionen*

$$v_t \colon \underline{\overline{\mathsf{X}}}_{-t} \to \mathbb{R} \qquad t \in \mathsf{T}_0$$

gibt, die folgendes leistet:

$$\begin{aligned} v_t(\xi) &= \operatorname*{val}_{\eta \in \underline{\overline{\mathsf{Y}}}_{-(t-1)}} \left\{ v_{t-1}(a_{-(t-1)}(\xi, \eta)) + f_{-(t-1)}(\xi, \eta) \right\} \qquad (1.6) \\ &= \operatorname*{val}_{\eta \in \underline{\overline{\mathsf{Y}}}_{-(t-1)}} \{ (\mathfrak{A}_t^\eta v_{t-1})(\xi) \} \qquad \xi \in \underline{\overline{\mathsf{X}}}_{-t}, t \in \mathsf{T} \\ v_0(\xi) &= u(\xi) \qquad \xi \in \underline{\overline{\mathsf{X}}}_T \qquad (1.7) \end{aligned}$$

$$\bar{\alpha}_t(\xi) \in \underline{\overline{\mathsf{Y}}}_{-(t-1)} \textit{ ist ein Paar von opt. Strategien in} \qquad (1.8)$$
$$\Gamma_{-(t-1)} = (\underline{\overline{\mathsf{Y}}}_{-(t-1)}^1, \underline{\overline{\mathsf{Y}}}_{-(t-1)}^2, (\mathfrak{A}_t^\bullet v_{t-1})(\xi))$$

Beweis. Wir wissen, daß $v_{t-1}(\xi)$ die Auszahlung an Spieler 1 im „one shot game“ $\Gamma_{-(t-1)}$ ist. Da ein Nullsummenspiel vorliegt, sind die Auszahlungen in jedem Gleichgewicht gleich dem Spielwert; die Gleichgewichte sind optimale Strategien. □

27. Mai 04

Beispiel 1.3.8. Nullsummenspiel, „stationär“, d. h.

$$\underline{\overline{\mathsf{X}}}_t = \ldots = \underline{\overline{\mathsf{X}}}_0 = \{\xi, \xi'\} \qquad \underline{\overline{\mathsf{Y}}}_t^i = \ldots = \underline{\overline{\mathsf{Y}}}_1^i = \{1,2,3\} \; i = 1,2$$

$$a_t = \ldots = a_1$$

$$a_1(\xi, \bullet) = \begin{pmatrix} \xi & \xi' & \xi' \\ \xi' & \xi & \xi' \\ \xi' & \xi' & \xi \end{pmatrix} \qquad f_1(\xi, \bullet) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned}
(\mathfrak{A}_1^\bullet w)(\xi) &= w(a_1(\xi)) + f_1(\xi, \bullet) \\
&= \begin{pmatrix} w(\xi) & w(\xi') & w(\xi') \\ w(\xi') & w(\xi) & w(\xi') \\ w(\xi') & w(\xi') & w(\xi) \end{pmatrix} \\
&\quad + \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} w(\xi)+1 & w(\xi')+1 & w(\xi')+1 \\ w(\xi') & w(\xi)+1 & w(\xi')+1 \\ w(\xi') & w(\xi') & w(\xi)+1 \end{pmatrix}
\end{aligned}$$

Ist $w(\xi') \geq w(x) \geq w(\xi') - 1$, so ist

$$\operatorname*{val}_{\underline{\overline{\mathsf{Y}}}_1}(\quad) = w(\xi) + 1 \quad \text{und} \quad \bar{\alpha}_1(\xi) = (1,1)$$

Lemma 1.3.9. *Sei Σ ein dynamisches Spiel und α ein Strategien-n-Tupel. Dann gilt für alle $\xi \in \underline{\overline{\mathsf{X}}}_{t-1}$*

$$C_\alpha^{i\xi} = C_\alpha^{i a_t^\alpha(\xi)} + f_t^i(\xi, \alpha_t(\xi))$$

Beweis. Sei $(X,Y) = (X^\alpha, Y^\alpha)$ der von α erzeugte Pfad, beginnend bei $\xi \in \underline{\overline{\mathsf{X}}}_{t-1}$. Dann haben wir

$$\begin{aligned}
C_\alpha^{i\xi} &= \sum_{s=t}^{T} f_s^i(x_{s-1}, y_s)) + u^i(x_T) \\
&= \sum_{s=t+1}^{T} f_s^i(x_{s-1}, y_s)) + u^i(x_T) + f_t^i(\xi, \alpha_t(\xi))
\end{aligned}$$

wegen $x_{t-1} = \xi$ und $y_t = \alpha_t(x_{t-1})$. Weil nun

$$\begin{aligned}
x_t &= a_t(x_{t-1}, y_t) \\
&= a_t(\xi, \alpha_t(\xi)) \\
&= a_t^\alpha(\xi)
\end{aligned}$$

ist, ist die Summe rechts Endauszahlung in der Tat gerade $C_\alpha^{i a_t(\xi)}$. □

Satz 1.3.10 (Zermelo-von Neumann-Kühn). *Sei Σ ein dynamisches Spiel und Γ_Σ^ξ die Familie der abgeleiteten Normalformen. Es sei $\bar{\alpha}$ durch Rückwärtsinduktion erhalten (teilspielperfekt!) sowie $(v_t^i)_{i \in I, t \in \mathsf{T}_0}$ die Familie der Wertfunktionen. Dann ist für alle t*

und ξ stets $\bar{\alpha}$ (oder genauer die Restriktion) ein Nash-Gleichgewicht in Γ_{Σ}^{ξ} und $v_{-t}^{i}(\xi)$ die Gleichgewichtsauszahlung., d. h.

$$C_{\bar{\alpha}}^{i\xi} = v_{-t}^{i}(\xi) \qquad \xi \in \underline{\overline{\mathsf{X}}}_t,\ i \in I$$

Beweis. **1. Schritt**
per Induktion:

$$\begin{aligned} v_t^i(\xi) &= v_{t-1}^i(a_{-(t-1)}^{\bar{\alpha}}(\xi)) + f^i(\xi, \bar{\alpha}_{-(t-1)}(\xi)) \\ &= C_{\bar{\alpha}}^{a_{-(t-1)}^{\bar{\alpha}}(\xi)} + f_{-(t-1)}^i(\xi, \bar{\alpha}_{-(t-1)}(\xi)) \\ &\text{nach Induktion, da } v_{t-1}^i = C_{\bar{\alpha}}^{i\xi} \text{ auf } \underline{\overline{\mathsf{X}}}_{-(t-1)} \\ &= C_{\bar{\alpha}}^{i\xi} \qquad \text{nach Lemma!} \end{aligned}$$

1. Juni 04

2. Schritt
Es weiche nun Spieler i von $\bar{\alpha}^i$ ab zu $\hat{\alpha}^i$. Setze $\hat{\alpha} := (\bar{\alpha}^1, \dots, \hat{\alpha}^i, \dots, \bar{\alpha}^n)$. Wir haben zu zeigen (induktiv):

$$C_{\bar{\alpha}}^{i\xi} \geq C_{\hat{\alpha}}^{i\xi}$$

Das wie folgt:

$$\begin{aligned} C_{\bar{\alpha}}^{i\xi} &= v_t^i(\xi) \\ &= v_{t-1}^i(a_{-(t-1)}^{\bar{\alpha}}(\xi)) + f_{-(t-1)}^i(\xi, \bar{\alpha}_{-(t-1)}(\xi)) \\ &\text{wie im ersten Schritt!} \\ &\geq v_{t-1}^i(a_{-(t-1)}^{\hat{\alpha}}(\xi)) + f_{-(t-1)}^i(\xi, \hat{\alpha}_{-(t-1)}(\xi)) \\ &\text{siehe rekursive One-Shot-Gleichung} \\ &= C_{\bar{\alpha}}^{ia_{-(t-1)}^{\hat{\alpha}}(\xi)} + f_{-(t-1)}^i(\xi, \hat{\alpha}_{-(t-1)}(\xi)) \\ &\text{nach erstem Schritt,} \quad C_{\alpha}^{i\xi} = v_t^i(\xi) \\ &\geq C_{\hat{\alpha}}^{ia_{-(t-1)}^{\hat{\alpha}}(\xi)} + f_{-(t-1)}^i(\xi, \hat{\alpha}_{-(t-1)}(\xi)) \\ &\text{per Induktion} \\ &= C_{\hat{\alpha}}^{i\xi} \qquad \text{nach Lemma für} \quad \hat{\alpha} \end{aligned}$$

□

Wir schließen noch an den „Test auf Gleichgewicht“.

Definition 1.3.11. *Eine Strategie $\bar{\alpha}$ erfüllt „one deviation principle“, falls für alle $i \in I$, alle $t \in \mathsf{T}$, alle $\xi \in \underline{\overline{\mathsf{X}}}_t$ und jedes $\hat{\alpha}^i$ mit*

$$\hat{\alpha}_s^i(\xi') = \bar{\alpha}_s^i(\xi') \qquad (s, \xi') \neq (t, \xi)$$

stets gilt:

$$C_{\bar{\alpha}}^{i\xi} \geq C_{\hat{\alpha}^i\bar{\alpha}^{-i}}^{i\xi}$$

Satz 1.3.12. *$\bar{\alpha}$ ist teilspielperfekt genau dann, wenn $\bar{\alpha}$ das „one deviation principle“ erfüllt. (D. h., man muß nur gegen einfache Abweichnungen testen!)*

Beweis. Klar: jedes teilapielperfekte Gleichgewicht erfüllt das „one deviation principle“ Umgekehrt: Durch Induktion. □

Beispiel 1.3.13. Wir betrachten zunächst ein dynamisches Nullsummenspiel. Es gibt 2 Spieler und 2 Zustände.

Zustand 1 Spieler 1 kontrolliert den Übergang (von 1 nach 2)

Zustand 2 Spieler 2 kontrolliert den Übergang (von 2 nach 1)

d. h.

$$\underline{\overline{\mathsf{X}}}_0 = \underline{\overline{\mathsf{X}}}_1 = \cdots = \underline{\overline{\mathsf{X}}}_t = \{1, 2\}.$$

Jeder Spieler entscheidet „für“ oder „gegen“ Wechsel, d. h.

$$\underline{\overline{\mathsf{Y}}}_1^1 = \underline{\overline{\mathsf{Y}}}_1^2 = \cdots = \underline{\overline{\mathsf{Y}}}_t^2 = \{+, -\}$$

Daher ist der Übergang beschrieben durch $a_0 = a_1 = \cdots = a_t$ mit

$$a_0(1, \bullet, \bullet) = \begin{array}{c c} & \begin{array}{cc} + & - \end{array} \\ \begin{array}{c} + \\ - \end{array} & \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \end{array}$$

$$a_0(2, \bullet, \bullet) = \begin{array}{c c} & \begin{array}{cc} + & - \end{array} \\ \begin{array}{c} + \\ - \end{array} & \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \end{array}$$

Die Kontrolle kostet etwas. Da ein Nullsummenspiel vorliegt, muß nur

$$f_0(\xi, \bullet, \bullet) = f_1(\xi, \bullet, \bullet) = \dots$$

angegeben werden. Es ist

$$f_0(1, \bullet, \bullet) = \begin{array}{c c} & \begin{array}{cc} + & - \end{array} \\ \begin{array}{c} + \\ - \end{array} & \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \end{array}$$

$$f_0(2, \bullet, \bullet) = \begin{array}{c c} & \begin{array}{cc} + & - \end{array} \\ \begin{array}{c} + \\ - \end{array} & \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix} \end{array}$$

Damit ist Σ erklärt. Rückwärtsinduktion liefert: $v_0(1) = u(1) = 6, v_0(2) = u(2) = 0$ sowie

$$\begin{aligned} v_t(1) &= \operatorname*{val}_{\underline{\overline{\mathsf{Y}}}_1^1 \times \underline{\overline{\mathsf{Y}}}_1^2} v_{t-1}(a_0(1, \bullet, \bullet) + f_0(1, \bullet, \bullet)) \\ &= \operatorname{val}\left\{ \begin{pmatrix} v_{t-1}(2) & v_{t-1}(2) \\ v_{t-1}(1) & v_{t-1}(1) \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \right\} \\ v_t(2) &= \operatorname{val}\left\{ \begin{pmatrix} v_{t-1}(1) & v_{t-1}(2) \\ v_{t-1}(1) & v_{t-1}(2) \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix} \right\} \end{aligned}$$

Insbesondere also

$$v_1(1) = \mathrm{val}\left\{\begin{pmatrix} 0 & 0 \\ 6 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 0 & 1 \\ 5 & 8 \end{pmatrix} = 5$$
$$v_1(2) = \mathrm{val}\left\{\begin{pmatrix} 6 & 0 \\ 6 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 6 & -1 \\ 8 & 1 \end{pmatrix} = 1$$

danach

$$\bar{\alpha}_T^1(1) = -\bar{\alpha}_T^2(1) = +$$

d. h. $\bar{\alpha}_T(1) = (-,+)$, $\bar{\alpha}_T(2) = (-,-)$

$$v_2(1) = \mathrm{val}\left\{\begin{pmatrix} 1 & 1 \\ 5 & 5 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 1 & 2 \\ 4 & 7 \end{pmatrix} = 4$$
$$v_2(2) = \mathrm{val}\left\{\begin{pmatrix} 5 & 1 \\ 5 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 5 & 0 \\ 7 & 2 \end{pmatrix} = 2$$

sowie

$$\bar{\alpha}_{T-1}(1) = (-,+) \qquad \bar{\alpha}_{T-1}(2) = (-,-)$$

und weiter

$$v_3(1) = \mathrm{val}\left\{\begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix} = 3$$
$$v_3(2) = \mathrm{val}\left\{\begin{pmatrix} 4 & 2 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 4 & 1 \\ 6 & 3 \end{pmatrix} = 3$$

sowie

$$\bar{\alpha}_{T-2}(1) = (-,+) \qquad \bar{\alpha}_{T-2}(2) = (-,-)$$

und weiter

$$v_4(1) = \mathrm{val}\left\{\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix} = 3$$
$$v_4(2) = \mathrm{val}\left\{\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}\right\} = \mathrm{val}\begin{pmatrix} 3 & 2 \\ 5 & 4 \end{pmatrix} = 4$$

sowie

$$\bar{\alpha}_{T-3}(1) = (+,+) \qquad \bar{\alpha}_{T-3}(2) = (-,-)$$

und weiter

$$v_5(1) = \mathrm{val}\begin{pmatrix} 4 & 5 \\ 2 & 5 \end{pmatrix} = 4$$
$$v_5(2) = \mathrm{val}\begin{pmatrix} 3 & 3 \\ 5 & 5 \end{pmatrix} = 5$$

sowie

$$\bar{\alpha}_{T-4}(1) = (+,+) \qquad \bar{\alpha}_{T-4}(2) = (-,-)$$

und weiter

$$v_5(1) = \mathrm{val}\begin{pmatrix} 5 & 6 \\ 3 & 6 \end{pmatrix} = 5$$
$$v_5(2) = \mathrm{val}\begin{pmatrix} 4 & 4 \\ 6 & 6 \end{pmatrix} = 6$$

sowie

$$\bar{\alpha}_{T-5}(1) = (+,+) \qquad \bar{\alpha}_{T-5}(2) = (-,-)$$

also: $v_t(1) = t-1$, $v_t(2) = t$, $\bar{\alpha}_{T-t+1}(1) = (+,+)$, $\bar{\alpha}_{T-t+1}(2) = (-,-)$

Beispiel 1.3.14. Maschinenerneuerung: Eine Maschine verursacht mit wachsendem Alter höhere Kosten. Ersatz ist ebenfalls kostspielig. Wann erneuern, was kostets?
Wir setzen

$$\underline{\overline{\mathsf{X}}}_t = \underline{\overline{\mathsf{X}}}_0 = [0,T] \qquad \underline{\overline{\mathsf{Y}}}_t = \underline{\overline{\mathsf{Y}}}_0 = \{0,1\}$$
$$\eta = 0 \text{ kein Neukauf} \qquad \eta = 1 \text{ Maschine ersetzen}$$

Dann ist

$$x_t = x_{t-1}(1-y_t) + 1$$

D. h.

$$a_t(\xi,\eta) = \xi(1-\eta) + 1 = a_1(\xi,\eta)$$

Betriebskosten und Ersatzkosten sind zeitabhängig/altersabhängig. Zu Periode t

$$\underbrace{y_t p_t(x_{t-1})}_{\text{Ersatzkosten}} + \underbrace{q_t(x_t)}_{\text{Betriebskosten}}$$

mit gewissen Funktionen $p_t(\bullet)$, $q_t(\bullet)$, d. h.

$$f_t(\xi,\eta) = \eta p_t(\xi) + q_t(\xi(1-\eta)+1)$$

Es handelt sich um 1-Personen-Min.-Problem. Die Opt.-Gleichung ist

$$\begin{aligned} v_t(\xi) &= \min_{\eta\in\{0,1\}} \{v_{t-1}(a_{-(t-1)}(\xi,\eta)) + f_{-(t-1)}(\xi,\eta)\} \\ &= \min_{\eta\in\{0,1\}} \{v_{t-1}(\xi(1-\eta)+1) + \eta p_{T-t+1}(\xi) + q(\xi(1-\eta)+1)\} \\ v_0(\xi) &= u(\xi) \end{aligned}$$

Wir versuchen:

$$\begin{aligned} p_0(\xi) &= p_t(\xi) = 20 \text{ TEURo} \\ q(\xi) &= \xi \\ u(\xi) &= \xi - 10 \quad \text{(Erlös bei Verkauf sind neg. Kosten)} \qquad T = 10 \end{aligned}$$

D. h.

$$a_t(\xi,\eta) = \xi(1-\eta)+1$$
$$f_t(\xi,\eta) = 20\eta + \xi(1-\eta)+1$$

und daher

$$\begin{aligned}
v_t(\xi) &= \min_{\eta\in\{0,1\}}\{v_{t-1}(\xi(1-\eta)+1)+20\eta+\xi(1-\eta)+1\}\\
v_0(\xi) &= \xi - 10 \quad \text{daher}\\
v_1(\xi) &= \min_{\eta\in\{0,1\}}\{\xi(1-\eta)+1-10+20\eta+\xi(1-\eta)+1\}\\
&= 2\xi - 8 + \min\{(20-2\xi)\eta\}\\
&= 2\xi - 8, \quad \text{da } \xi \le 10
\end{aligned}$$

und $\alpha_T(\xi) = 0\ \forall \xi \in \underline{\mathrm{X}}_0$. Für die vorletzte Entscheidung hat man

$$\begin{aligned}
v_2(\xi) &= \min_{\eta\in\{0,1\}}\{2(\xi(1-\eta)+1)-8+20\eta+\xi(1-\eta)+1\}\\
&= 3\xi - 5 + \min_{\eta\in\{0,1\}}\{20-3\xi)\eta\}\\
&= \begin{cases} 3\xi-5 & \xi\le 6\\ 15 & \xi \ge 7\end{cases}\\
\text{oder}\quad v_2(\xi) &= \min\{3\xi-5, 15\}\\
\text{sowie}\quad \bar{\alpha}_{T-1}(\xi) &= \begin{cases} 0 & \xi \le 6\\ 1 & \xi\ge 7\end{cases}
\end{aligned}$$

Im nächsten Schritt:

$$v_3(\xi) = \min_{\eta\in\{0,1\}}\{v_2(\xi(1-\eta)+1)+20\eta+\xi(1-\eta)+1\}$$

1) Ist $\xi \le 5$, so ist $\xi(1-\eta))+1 \le \xi+1 \le 6$, also ist $v_2(\bullet) = 3\bullet -5$ zu nehmen, d. h.

$$\begin{aligned}
v_3(\xi) &= \min_{\eta\in\{01,\}}\{3(\xi(1-\eta)+1)-5+20\eta+\xi(1-\eta)+1\}\\
&= 4\xi - 1 + \min_{\eta\in\{0,1\}}\{(20-4\xi)\eta\}\\
&= 4\xi - 1 \quad \forall \xi \le 5
\end{aligned}$$

$$\text{sowie}\quad \bar{\alpha}_{T-2}(\xi) = 0 \quad \forall \le 5$$

2) für $\xi \ge 6$ ist

$$\begin{aligned}
v_2(\xi(1-\eta)+1) &= \begin{cases} v_2(1) & \eta = 1\\ v_2(\xi+1) & \eta = 0\end{cases} = \begin{cases} -2 & \eta=1\\ 15 & \eta = 1\end{cases}\\
\text{d. h.}\quad v_3(\xi) &= \min_\eta\{-2+20+1, 15+\xi+1\}\\
&= \min_\eta\{19, 16+\xi\} \quad \forall \xi \ge 6\\
&= 19\\
\text{sowie}\quad \bar{\alpha}_{T-2}(\xi) &= 1
\end{aligned}$$

2 Varianten: Stoppen / Diskontierung

2.1 Stoppzeiten

Gelegentlich sind Prozesse von verschiedener Dauer, Laufzeit hängt vom Pfad ab (Zusammenbruch einer Maschine, Lager kann Nachfrage nicht befriedigen) Im Prinzip wird eine „pfadabhängige Zeit“

$$\tau \colon \underline{\overline{\mathsf{X}}} \to T$$

konstruiert und die Auszahlung ist

$$\sum_{t=0}^{\tau} f_t^i(x_{t-1}, y_t) + u^i(x_\tau)$$

mit $x = x^{\alpha_i \xi}$, $y = y^{\alpha_i \xi}$ und daher auch $\tau(x^{\alpha_i \xi}) = \tau$.
Problem: Stoppzeit $\tau(x)$ sollte nicht von x_t $(t > \tau)$ abhängen! Etwa: sei

$$\tau \colon \underline{\overline{\mathsf{X}}} \to \mathsf{T} \qquad \text{gegeben durch}$$
$$\tau(x) = \begin{cases} 8 & x_{10} = 0 \\ 10 & x_{10} \neq 0 \end{cases}$$

„Stoppen bei 8, wenn der Pfad bei 10 die Null erreicht.“ Man könnte bei $t = 8$ nicht entscheiden, ob zu stoppen ist... Betrachte für vorgegebenes τ die Menge

$$S_t = \{x' \mid \tau(x') = t\}$$

für festes t. Die Entscheidung $x \in S_t$ sollte nur von $x_1, \ldots, x_t$ abhängen. D. h. die Indikatorfunktion

$$1_{S_t} \colon \underline{\overline{\mathsf{X}}} \to \{0, 1\}$$
$$1_{S_t}(x) = \begin{cases} 0 & x \notin S_t \\ 1 & x \in S_t \end{cases}$$

sollte nur von $x_0, \ldots, x_t$ abhängen.

Definition 2.1.1. *Sei* $\Sigma = (\mathfrak{X}, \mathfrak{Y}, a, f, u, T)$ *ein dynamisches Spiel / Entscheidungsproblem. Eine* Stoppzeit *ist eine Abbildung*

$$\tau \colon \underline{\overline{\mathsf{X}}} \to \mathsf{T}$$

derart, daß gilt: Für jedes $t \in \mathsf{T}$ $\exists$ *Funktion*

$$h_t \colon \prod_{s=0}^{t} \underline{\overline{\mathsf{X}}}_t \to \{0, 1\},$$

die $1_{S_t}(x) = h_t(x_0, \ldots, x_t)$ $\forall x \in \underline{\overline{\mathsf{X}}}$ *leistet.*

Beispiel 2.1.2 *(„Zeit des ersten Eintritts“, „First entrance time“).* Sei $\underline{\overline{X}}_0 = \underline{\overline{X}}_1 = \ldots = \underline{\overline{X}}_T$ sowie $F \subseteq \underline{\overline{X}}_0$. Wir setzen $\tau(x) := \min\{t \geq 1 \mid x_t \in F\}$ („Stoppe, sobald x erstmals in F eintritt“) τ ist Stoppzeit: Es ist

$$\begin{aligned} S_t &= \{x \mid \tau(x) = t\} \\ &= \{x \mid \min\{s \mid x_s \in F\} = t\} \\ &= \{x \mid x_s \notin F \ \forall s < t, x_t \in F\} \\ &= \bigcap_{s=1}^{t-1} \{x \mid x_s \notin F\} \cap \{x \mid x_t \in F\} \end{aligned}$$

Wegen $1_{F \cap G} = 1_F \wedge 1_G$ ($\wedge$ sei Min.) ist

$$1_{S_t} = \bigwedge_{s=1}^{t} 1_{\{x' \mid x'_s \notin F\}} \wedge 1_{\{x' \mid x'_t \in F\}}$$

15. Juni reicht es zu zeigen: $h_t(x) = 1_{\{x' \mid x'_s \in F\}}(x)$ eine Funktion von $x_0, \ldots, x_t$ ist (für $s \leq t$)

Lemma 2.1.3. *Genau dann ist τ eine Stoppzeit, wenn es Mengen $H_t \subseteq \underline{\overline{X}}_0 \times \cdots \times \underline{\overline{X}}_t$ gibt, so daß $1_{S_t}(x) = 1_{H_t}(x_0, \ldots, x_t)$ gilt für $t \in \mathsf{T}$.*

Beweis. Klar, man muß nur mit der vorhergehenden Definition

$$H_t = \{(x_0, \ldots, x_t) \mid h_t(x_0, \ldots, x_t) = 1\}$$

setzen. □

Weitere

Beispiele 2.1.4.

1) $\underline{\overline{X}}_0 = \ldots = \underline{\overline{X}}_t$, $F \subseteq \underline{\overline{X}}_0$. Setze $\tau = \min\{t \mid x_t \in F$ zum zweiten Mal$\}$. Dann folgt $\tau = \min\{t \mid \exists t' < t \ : \ x_{t'}, x_t \in F, x_s \notin F \forall s \neq t, s \neq t'\}$

2) Sei $\bar{x} \in \underline{\overline{X}}$ fixiert und $\tau(x) = \min\{t \mid |x_t - \bar{x}_t| \leq 10\}$ (für $\underline{\overline{X}}_t \subseteq \mathbb{R}, \mathbb{N}, \ldots$, d. h., $|\bullet|$ existiert). D. h. τ ist eine verallgemeinerte first-entrance-Zeit: es gibt Mengen $F_t \subseteq \underline{\overline{X}}_t$ $\forall t = 1, \ldots, T$ so daß

$$\tau(x) = \min\{t \mid x_t \in F_t\}$$

gilt. Nämlich hier: $F_t = \{\xi \mid |\xi - \bar{x}_t| \leq 10\}$

Satz 2.1.5.

1) *τ ist Stoppzeit genau dann, wenn mit $R_t = \{x \mid \tau(x) \leq t\}$ stets 1_{R_t} nur von $(x_1, \ldots, x_t)$ abhängt.*

2) *Sind τ, σ Stoppzeiten, so auch $\tau \wedge \sigma$ und $\tau \vee \sigma$ (Min., Max. von τ und σ).*

3) *Konstante $\tau = t$ sind Stoppzeiten.*

Beweis. 1) Es ist

$$\begin{aligned} R_t &= \bigcup_{s=1}^{t} \{x \mid \tau(x) = s\} \\ &= \bigcup_{s=1}^{t} S_s \qquad \text{und daher} \end{aligned}$$

$$1_{R_t} = 1_{\bigcup_{s=1}^t S_s} = \bigvee_{s=1}^{t} 1_{S_s}$$

2) $\{x \mid \tau(x) \wedge \sigma(x) \leq t\} = \{x \mid \tau(x) \leq t\} \cup \{x \mid \sigma(x) \leq t\}$

3) Klar.

$\square$

Für die dynamische Optimierung sollen nur verallgemeinerte f.-e.-Zeiten benutzt werden. D. h. zu jedem t existiert eine Menge $\partial\underline{\overline{\mathsf{X}}}_t$ („Rand von $\underline{\overline{\mathsf{X}}}_t$“) mit $\tau(x) = \min\{t \mid x_t \in \partial\underline{\overline{\mathsf{X}}}_t\}$. Wir schreiben auch: $\overset{\circ}{\underline{\overline{\mathsf{X}}}}_t := \underline{\overline{\mathsf{X}}}_t - \partial\underline{\overline{\mathsf{X}}}_t \ \forall t \in \mathsf{T}$. Dann hat man

Definition 2.1.6. *Ein dynamisches Spiel mit Stoppzeit ist ein Tupel*

$$\Sigma = (\mathfrak{X}, \mathfrak{Y}, a, f, u, \tau)$$

Dabei ist $\mathfrak{X} = (\underline{\overline{\mathsf{X}}}_0, \dots, \underline{\overline{\mathsf{X}}}_t)$, $\mathfrak{Y}$ *wie bisher,* τ *verallgemeinerte f. e.-Zeit (also* $\overset{\circ}{\underline{\overline{\mathsf{X}}}}_t, \partial\underline{\overline{\mathsf{X}}}_t$ *sind definiert),*

$$\begin{aligned} a_t &\colon \overset{\circ}{\underline{\overline{\mathsf{X}}}}_{t-1} \times \underline{\overline{\mathsf{Y}}}_t \to \underline{\overline{\mathsf{X}}}_t, \\ f_t^i &\colon \overset{\circ}{\underline{\overline{\mathsf{X}}}}_{t-1} \times \underline{\overline{\mathsf{Y}}}_t \to \mathbb{R} \qquad \textit{sowie} \\ u &= (u_t^i)_{t \in \mathsf{T}}^{i \in I} \qquad \textit{mit} \\ u_t^i &\colon \partial\underline{\overline{\mathsf{X}}}_t \to \mathbb{R} \end{aligned}$$

Die relevanten Pfade sind von der Form $x = (x_0, \dots, x_s)$ *mit* $x_0, \dots, x_{s-1} \in \overset{\circ}{\underline{\overline{\mathsf{X}}}}_{0,\dots,s-1}$, $x_s \in \partial\underline{\overline{\mathsf{X}}}_s$, *so daß* $\tau(x) = s$ *ist.*
Strategien: $\alpha_t^i \colon \overset{\circ}{\underline{\overline{\mathsf{X}}}}_{t-1} \to \underline{\overline{\mathsf{Y}}}_t^i \ \forall i \in I, t \in \mathsf{T}$; *Auszahlungen sind wie üblich:*

$$C_\alpha^{i\xi} = \sum_{t=1}^{\tau} f_t^i(x_{t-1}, y_t) + u_\tau^i(x_\tau) \text{ für } \xi \in \underline{\overline{\mathsf{X}}}_0,$$

ebenso für $\xi \in \overset{\circ}{\underline{\overline{\mathsf{X}}}}_t$. *Dabei ist* $X = x^{\alpha\xi}$ *von* ξ *und* α *abhängig. daher ist auch* $\tau = \tau(x^{\xi\alpha})$ *von* ξ, α *abhängig. Damit ist die Normalform*

$$\Gamma = \Gamma_\Sigma^\xi = (\sigma^1, \dots, \sigma^n, C_\bullet^{1\xi}, \dots, C_\bullet^{n\xi})$$

Der weitere Fortgang wie üblich:

Definition 2.1.7. *Man sagt,* $\bar{\alpha}$ *sei teilspielperfekt, oder durch Rückwärtsinduktion erhalten, falls es eine Familie*

$$v_t^i \colon \underline{\overline{\mathsf{X}}}_{-t} \to \mathbb{R} \qquad \forall i \in I, t \in \mathsf{T}$$

gibt mit

$$\begin{aligned} v_t^i(\xi) &= \max_{\eta^i \in \underline{\overline{\mathsf{Y}}}^i_{-(t-1)}} \{(\mathfrak{A}_t^{i\eta^i\bar{\alpha}} v_{t-1}^i)(\xi)\} && (*) \\ &= (\mathfrak{A}_t^{i\bar{\alpha}} v_{t-1}^i)(\xi) \qquad \forall \xi \in \overset{\circ}{\underline{\overline{\mathsf{X}}}}_{-t}, t \in \mathsf{T} \\ v_t^i(\xi) &= u_t^i(\xi) \qquad \forall \xi \in \partial\underline{\overline{\mathsf{X}}}_{-t} && (**) \end{aligned}$$

mit $a_t^\alpha(\xi) = a_t(\xi, \alpha_t(\xi)), (\mathfrak{A}_t^{i\alpha} w)(\xi) = w(a_t^\alpha(\xi)) + f^i(\xi, a_t^\alpha(\xi))$

Natürlich gilt der

Satz 2.1.8. *Sei Σ ein dynamisches Spiel mit Stoppzeit. Sei $\bar{\alpha}$ teilspielperfekt und v_t^i die Familie der Wertfunktion. Dann ist für alle t und $\xi \in \overset{\circ}{\underline{\mathsf{X}}}_t$ stets $\bar{\alpha}$ ein Nash-Gleichgewicht in Γ_Σ^ξ und $v_t^i(\xi)$ die Gleichgewichts-Auszahlung.*

17. Juni

2.2 Unendlicher Horizont, Diskontierung

Wir betrachten ein stationäres Modell mit

$$\begin{aligned} \underline{\mathsf{X}}_0 &= \underline{\mathsf{X}}_1 = \dots & \qquad \underline{\mathsf{Y}}_1 &= \underline{\mathsf{Y}}_2 = \dots \\ a_1 &= a_2 = \dots & f_1 &= f_2 = \dots \end{aligned}$$

Wir beschränken uns auf Nullsummenspiele (gilt auch für min/max-Problem). Die Optimalitätsgleichung lautet dann

$$v_t(\xi) = \operatorname*{val}_{\eta \in \underline{\mathsf{Y}}_1} \{v_{t-1}(a_1(\xi, \eta)) + f_1(\xi, \eta)\}$$

Wir erinnern uns an den induzierten Operator: Ist $\mathcal{F}(\underline{\mathsf{X}}_0) = \{w \mid w \colon \underline{\mathsf{X}}_0 \to \mathbb{R}\}$, so ist

$$\begin{aligned} &\mathfrak{A}^\eta \colon \mathcal{F}(\underline{\mathsf{X}}_0) \to \mathcal{F}(\underline{\mathsf{X}}_0) \\ &\quad (\mathfrak{A}^\eta w)(\eta) := w(a_1(\xi, \eta)) + f_1(\xi, \eta) \quad \xi \in \underline{\mathsf{X}}_0, w \in \mathcal{F}(\underline{\mathsf{X}}_0) \end{aligned}$$

Solange $\underline{\mathsf{X}}_0$ endlich ist, ist

$$\mathcal{F}(\underline{\mathsf{X}}_0) = \mathbb{R}^{|\underline{\mathsf{X}}_0|}$$

Vorausüberlegungen zu ∞-Horizont: Wir wollen $T \to \infty$ durchführen. Jedoch würde

$$u(x_T) + \sum_{t=1}^{T} f_0(x_{t-1}, \alpha(x_{t-1}))$$

nicht konvergieren. Ökonomisch sinnvoll ist Einführung eines *Diskontfaktors:* naiv: ein Kapital K ist nach einem Jahr $K(1 + p/100)$ wert. Umgekehrt ist ein Kapital K, das erst in einem Jahr verfügbar wird, jetzt $K/(1 + p/100) =: K \cdot \rho$ mit $0 < \rho < 1$ wert, d. h., K wird *abdiskontiert.* Wir ändern daher die Auszahlung zu

$$C_\alpha^\xi = \rho^T u(x_T) + \sum_{t=1}^{T} \rho^{t-1} f(x_{t-1}, \alpha_t(x_{t-1}))$$

Da f auf endl. $\underline{\mathsf{X}}_0, \underline{\mathsf{X}}_1$ beschränkt ist, konvergiert alles stets.

$$C_\alpha^\xi \leq c \sum_{t=1}^{\infty} \rho^{t-1} \qquad c \text{ const.}$$

Dies erinnert wieder an einen nicht stationären Fall (mit $f_t^* = \rho 1t - 1f$). Die Optimalitätsgleichung für Horizont T würde nun lauten

$$v_t(\xi) = \operatorname*{val}_{\eta \in \underline{\mathsf{Y}}_1} \{v_{t-1}(a(\xi, \eta)) + \rho^{T-t} f(\xi, \eta)\} \tag{$*$}$$

Wir streben die Definition einer Wertfunktion v für ∞-Horizont an; unabhängig von t. Beginnt man einen Schritt später, so sieht das ∞-Horizont-Problem unverändert aus. Der Zusammenhang mit T-Horizont sollte sein

$$v_t = \rho^{T-t} v.$$

Einsetzen in $(*)$ liefert

$$\begin{aligned} \rho^{T-t} v(\xi) &= \operatorname*{val}_{\eta \in \underline{\overline{\mathsf{Y}}}_1} \{\rho^{T-(t-1)} v(a(\xi,\eta)) + \rho^{T-t} f(\xi,\eta)\} \\ \text{d. h.} \quad v(\xi) &= \operatorname*{val}_{\eta \in \underline{\overline{\mathsf{Y}}}_1} \{\rho v(a(\xi,\eta)) + f(\xi,\eta)\} \end{aligned}$$

Dies ist auch direkt zu interpretieren! Zu klären:

1) Existiert v?

2) Kann v als Spielwert hinsichtlich der Auszahlung

$$\lim_T \rho^T u(x_t) + \sum_{t=1}^{T} \rho^{t-1} f(x_{t-1}, \alpha_t(x_{t-1}))$$

erkannt werden?

3) Für u beliebig? Ohne u?

4) Gibt es auch stationäre (oder nicht stationäre) opt. Strategien?

Definition 2.2.1. *Ein stationäres dynamisches Spiel mit Diskontierung ist ein Tupel*

$$\Sigma = (\mathfrak{X}, \mathfrak{Y}, a, f, \rho)$$

mit

$$\begin{aligned} \mathfrak{X} &= (\underline{\overline{\mathsf{X}}}_0, \underline{\overline{\mathsf{X}}}_0, \dots) \\ \mathfrak{Y} &= (\underline{\overline{\mathsf{Y}}}_1, \underline{\overline{\mathsf{Y}}}_1, \dots) \\ 0 < \rho &< 1 \end{aligned}$$

Eine Strategie (für Spieler i) ist $\alpha_t^i \colon \underline{\overline{\mathsf{X}}}_0 \to \underline{\overline{\mathsf{Y}}}_1$; α^i ist stationär, falls

$$\alpha^i = (\alpha_1^i, \alpha_1^i, \dots)$$

Die Auszahlung ist

$$C_\alpha^{i\xi} = \sum_{t=1}^{\infty} \rho^{t-1} f(x_{t-1}, \alpha_t(x_{t-1}))$$

stets gleichmäßig konvergent, da dominiert von geometrischer Reihe). Also existiert die Normalform

$$\Gamma_\Sigma^\xi = (\mathfrak{S}^1, \mathfrak{S}^2, C_\bullet^\xi) \qquad \textit{für } i = 1$$

Definition 2.2.2. *Für* $w \in \mathcal{F}(\overline{\underline{X}}_0)$ *schreiben wir*

$$\|w\| := \max_{\xi \in \overline{\underline{X}}_0} |w(\xi)|$$

$\|\bullet\|$ *ist die* ∞*-Norm auf* $\mathcal{F}(\overline{\underline{X}}_0)$*. Für* $\eta \in \overline{\underline{Y}}_1$ *sei*

$$\begin{aligned} \mathfrak{A}^\eta \colon \mathcal{F}(\overline{\underline{X}}_0) &\to \mathcal{F}(\overline{\underline{X}}_0) \\ (\mathfrak{A}^\eta w)(\xi) &= \rho w(a(\xi,\eta)) + f(\xi,\eta) \end{aligned}$$

Ferner betrachten wir zwei Versionen eines Operators mit val*:*

$$(\mathfrak{A}w)(\xi) := \operatorname*{val}_{\eta \in \overline{\underline{Y}}_1} (\mathfrak{A}^\eta w)(\xi)$$

–wir würden dann nach v *mit* $v = \mathfrak{A}v$ *suchen. Da* val *oft nicht existiert, betrachten wir auch gemischte Strategien! Dazu sei*

$$\mathfrak{M}^i = \{a \in \mathbb{R}^{\overline{\underline{Y}}_1^i} \mid a \geq 0, \sum_{\eta \in \overline{\underline{Y}}_1^i} a_\eta = 1\}$$

die Menge der gemischten Strategien von i *(zu fester Zeit). Wir können dann definieren*

$$\begin{aligned} \mathfrak{A} \colon \mathcal{F}(\overline{\underline{X}}_0) &\to \mathcal{F}(\overline{\underline{X}}_0) \\ (\mathfrak{A}w)(\xi) &= \operatorname*{val}_{\mathfrak{M}^1 \times \mathfrak{M}^2} \{(\mathfrak{A}^\bullet w)(\xi)\}, \end{aligned}$$

dieser existiert stets.

Bemerkung 2.2.3. Für jedes $\eta \in \overline{\underline{Y}}_1$ gilt

$$\|\mathfrak{A}^\eta w - \mathfrak{A}^\eta w'\| \leq \rho \|w - w'\| \qquad \forall w, w' \in \mathcal{F}(\overline{\underline{X}}_0),$$

d. h., der Operator ist kontrahierend.

Beweis. Klar ist

$$\begin{aligned} (\mathfrak{A}^\eta w)(\xi) - (\mathfrak{A}^\eta w')(\xi) &= \rho w(a(\xi,\eta)) - \rho w'(a(\xi,\eta)) \\ \text{also} \quad \max_\xi \big|(\mathfrak{A}^\eta w)(\xi) - (\mathfrak{A}^\eta w')(\xi)\big| &= \rho \max_\xi \big|w(a(\xi,\eta)) - w'(a(\xi,\eta))\big| \\ &\leq \rho \max_{\xi^*} \big|w(\xi^*) - w'(\xi^*)\big| \\ &= \rho \|w - w'\| \end{aligned}$$

□

Satz 2.2.4. $\mathfrak{A}$ *ist strikt kontrahierend, d. h., für* $w, w' \in \mathcal{F}(\mathfrak{X}_0)$ *gilt*

$$\|\mathfrak{A}w - \mathfrak{A}w'\| \leq \rho \|w - w'\| < \|w - w'\|$$

Beweis. Wir zeigen zunächst: Seien C, C' zwei $\underline{\overline{Y}}_1^1 \times \underline{\overline{Y}}_1^2$-Matrizen. Dann gilt:

$$\left| \operatorname*{val}_{\mathfrak{M}^1 \times \mathfrak{M}^2} C - \operatorname*{val}_{\mathfrak{M}^1 \times \mathfrak{M}^2} C' \right| \leq \max_{\underline{\overline{Y}}_1^1 \times \underline{\overline{Y}}_1^2} \left| C_{\eta^1 \eta^2} - C'_{\eta^1 \eta^2} \right| \tag{$*$}$$

$$\operatorname{val} C = \operatorname{val} a^1 C a^2 \qquad (a^1, a^2) \in \mathfrak{M}^1 \times \mathfrak{M}^2$$

($\mathfrak{M}^1$: gemischte Strategien über $\underline{\overline{Y}}_1^1$). Das gilt wegen:

$$\begin{aligned}
\operatorname*{val}_{(a^1,a^2)} a^1 C a^2 - \operatorname*{val}_{(a^1,a^2)} a^1 C' a^2 &= \max_{a^1 \in \mathfrak{M}^1} \min_{a^2 \in \mathfrak{M}^2} a^1 C a^2 - \min_{a^2} \max_{a^1} a^1 C' a^2 \\
&= \min_{a^2 \in \mathfrak{M}^2} \bar{a}^1 C a^2 - \max_{a^1} a^1 C' \bar{a}^2 \qquad \text{für passende } (\bar{a}^1, \bar{a}^2) \\
&\leq \bar{a}^1 C \bar{a}^2 - \bar{a}^1 C' \bar{a}^2 \\
&= \bar{a}^1 (C - C') \bar{a}^2 \\
&= \sum_{\eta_1, \eta_2} \bar{a}^1_{\eta_1} (C_{\eta_1,\eta_2} - C'_{\eta_1,\eta_2}) \bar{a}^2_{\eta_2} \\
&\leq \max_{\eta} \left| C_\eta - C'_\eta \right|
\end{aligned}$$

Aus Symmetriegründen ist auch

$$\operatorname{val} C' - \operatorname{val} C \leq \max_{\eta} \left| C_\eta - C'_\eta \right|,$$

das zeigt ($*$). Nun wenden wir dies an auf

$$C_\eta = (\mathfrak{A}^\eta w)(\xi), \qquad C'_\eta = (\mathfrak{A}^\eta w')(\xi)$$

Es ist

$$\begin{aligned}
\|\mathfrak{A} w - \mathfrak{A} w'\| &= \max_{\xi} \left| (\mathfrak{A} w)(\xi) - (\mathfrak{A} w')(\xi) \right| \\
&= \max_{\xi} \left| \operatorname{val}(\mathfrak{A}^\eta w)(\xi) - \operatorname{val}(\mathfrak{A}^\eta w')(\xi) \right| \\
&\leq \max_{\xi} \left| \max_{\eta} \left| \mathfrak{A}^\eta w(\xi) - \mathfrak{A}^\eta w'(\xi) \right| \right| \\
&= \max_{\eta} \max_{\xi} \left| (\mathfrak{A}^\eta w)(\xi) - (\mathfrak{A}^\eta w')(\xi) \right| \\
&= \max_{\eta} \|\mathfrak{A}^\eta w - \mathfrak{A}^\eta w'\| \\
&\leq \rho \|w - w'\|
\end{aligned}$$

□

Satz 2.2.5. *Sei* $v_0 \colon \underline{\overline{X}}_0 \to \mathbb{R}$ *beliebig. sowie*

$$\begin{aligned}
v_1 &= \mathfrak{A} v_0 \\
v_2 &= \mathfrak{A} v_1 = \mathfrak{A}^2 v_0 \\
&\vdots \\
v_t &= \mathfrak{A} v_{t-1} = \mathfrak{A}^t v_0
\end{aligned}$$

Dann gilt:

1) $\exists\ \lim_{t\to\infty} \mathfrak{A}^t v = \lim_{t\to\infty} v_t =: v$ *gleichmäßig in* ξ *(d. h. in* $\|\bullet\|$*)*

2) Es gilt $v = \mathfrak{A}v$*, also* v *ist Fixpunkt.*

3) v *ist der einzige Fixpunkt von* $\mathfrak{A}$*.*

Beweis. **1. Schritt**
Wir beachten für natürliches s:

$$\begin{aligned} \|\mathfrak{A}^s v_0 - \mathfrak{A}^{s-1} v_0\| &= \|\mathfrak{A}\mathfrak{A}^{s-1} v_0 - \mathfrak{A}\mathfrak{A}^{s-2} v_0\| \\ &\le \rho \|\mathfrak{A}^{s-1} v_0 - \mathfrak{A}^{s-2} v_0\| \\ &\le \rho^{s-1} \|\mathfrak{A} v_0 - v_0\| \end{aligned}$$

Daher ist $\mathfrak{A}^t v$ eine Cauchy-Folge:

$$\begin{aligned} \|\mathfrak{A}^t v_0 - \mathfrak{A}^r v_0\| &= \| \sum_{s=r+1}^{t} \mathfrak{A}^s v_0 - \mathfrak{A}^{s-1} v_0\| \\ &\le \sum_{s=r+1}^{t} \|\mathfrak{A}^s v_0 - \mathfrak{A}^{s-1} v_0\| \\ &\le \sum_{s=r+1}^{t} \rho^{s-1} \|\mathfrak{A} v_0 - v_0\| \\ &\le \rho^r \text{ const.} \end{aligned}$$

Also existiert $v := \lim_{t\to\infty} v_t$, d. h.

$$\|v_t - v\| \xrightarrow[t\to\infty]{} 0$$

2. Schritt
Wir vermuten $\mathfrak{A}v = v$ wegen

$$\mathfrak{A}v = \mathfrak{A} \lim_t \mathfrak{A}^t v_0 \overset{?}{=} \lim_t \mathfrak{A}^{t+1} v_0 = v$$

genauer:

$$\begin{aligned} \|\mathfrak{A}v - v\| &= \|\mathfrak{A}v - \mathfrak{A}^{t+1} v_0 + \mathfrak{A}^{t+1} v_0 - v\| \\ &\le \|\mathfrak{A}v - \mathfrak{A}^{t+1} v_0\| + \|\mathfrak{A}^{t+1} v_0 - v\| \\ &\le \rho \underbrace{\|v - \mathfrak{A}^t v_0\|}_{\xrightarrow[t\to\infty]{} 0} + \underbrace{\|\mathfrak{A}^{t+1} v_0 - v\|}_{\xrightarrow[t\to\infty]{} 0} \end{aligned}$$

Also $\|\mathfrak{A}v - v\| = 0$, d. h. $\mathfrak{A}v = v$.

3. Schritt
v ist eindeutig: Wäre w ein zweiter Fixpunkt, so hätte man

$$\begin{aligned} \|v - w\| &= \|\mathfrak{A}v - \mathfrak{A}w\| \\ &\le \rho \|v - w\| \qquad 0 < \rho < 1 \end{aligned}$$

also

$$\begin{aligned}\|v-w\| &= 0\\ v &= w\end{aligned}$$

□

Satz 2.2.6. *Sei v der eindeutige Fixpunkt von $\mathfrak{A}$, d. h.,*

$$v(\xi) = \operatorname{val}(\mathfrak{A}^\eta v)(\xi) = \operatorname{val}\{\rho v(a(\xi,\eta)) + f(\xi,\eta)\} \tag{$*$}$$

Es sei für jedes $\xi \in \underline{\overline{X}}_0$ jeweils $\bar{\alpha}_0(\xi)$ ein Paar von reinen optimalen Strategien von $\Gamma^\xi = (\underline{\overline{Y}}_1, (\mathfrak{A}^\eta v)(\xi))$. Dann hat $\Gamma^\xi_\Sigma = (\mathfrak{S}^1, \mathfrak{S}^2, C^\xi_\bullet)$ Spielwert und opt. Strategien. Insbesondere ist $\bar{\alpha} = (\bar{\alpha}_0, \bar{\alpha}_0, \dots)$ und

$$v_{\Gamma^\xi_\Sigma} = v(\xi).$$

Beweis. **1. Schritt**
Wir betrachten nur das dyn. Spiel mit endlichem Horizont

$$\Sigma^{*T} = (\mathfrak{X}, \mathfrak{Y}, a, f^*, u^*, T)$$

mit $\mathfrak{X} = (\underbrace{\underline{\overline{X}}_0, \dots, \underline{\overline{X}}_0}_{T+1\text{-mal}})$ und $\mathfrak{Y} = (\underbrace{\underline{\overline{Y}}_0, \dots, \underline{\overline{Y}}_0}_{T\text{-mal}})$, a wie in Σ, $f^*_t = \rho^{t-1} f$, $u^* = \rho^T v$, wobei v der Fixpunkt von $\mathfrak{A}$ gemäß früherem Satz ist.
Rückwärtsinduktion in Σ^{*T} liefert nun

$$\begin{aligned}
v_0 &= u^* = \rho^T v\\
v_1(\bullet) &= \operatorname*{val}_{\underline{\overline{Y}}_1}\{f^*(\xi,\bullet) + u^*(a(\xi,\bullet))\} \qquad (**)\\
&= \operatorname{val}\{\rho^T v(a(\xi,\bullet)) + \rho^{T-1} f(\xi,\bullet)\}\\
&= \rho^{T-1}\operatorname{val}\{\rho v(a(\xi,\bullet)) + f(\xi,\bullet)\}\\
&= \rho^{T-1}\mathfrak{A}v = \rho^{T-1} v\\
v_2 &= \rho^{T-2} v\\
&\vdots\\
v_T &= v
\end{aligned}$$

D. h., der Fixpunkt v ist die Wertfunktion des endlichen Problems Σ^{*T} – unabhängig von T.

2. Schritt
Optimale Strategien in Σ^{*T} erhält man durch Auswahl von optimalen Strategien auf jeder Stufe, d. h., bei der Wertberechnung in $(**)$! Nun ist das val-Problem in $(**)$ beim Rückwärts-Rechnen in Σ^{*T} jedesmal gleich (bis auf ρ^{T-1}, ρ^{T-2} usw., aber das liefert gleiche optimale Strategien!) Wählt man also auf jeder Stufe stets $\bar{\alpha}_0$ als optimale Strategie, so ist

$$\bar{\alpha}^* = (\underbrace{\bar{\alpha}_0, \dots, \bar{\alpha}_0}_{T\text{-mal}})$$

eine optimale Strategie in $\Gamma_{\Sigma^{*T}}$.

3. Schritt
Ist α irgendeine Strategie für Σ (∞-Horizont, nicht notw. stationär!), so liefert α auch eine Strategie in Σ^{*T} (per Restriktion). Man sieht leicht:

$$x_s^{*T} = x_s = x_s^\alpha \qquad (s = 0, \dots, T)$$

(d. h. α erzeugt den gleichen Pfad bis T für Σ und Σ^{*T}) Daher ist

$$\begin{aligned} C_\alpha^\xi &= \sum_{t=1}^{\infty} \rho^{t-1} f(x_{t-1}, \alpha_t(x_{t-1})) \\ &= \sum_{t=1}^{T} \dots + \sum_{t=T+1}^{\infty} \dots \\ &= \sum_{t=1}^{T} \dots + \rho^T v(x_T) - \rho^T v(x_T) + \sum_{t=T+1}^{\infty} \dots \\ &= C_\alpha^{*T\xi} - \rho^T v(x_T) + \sum_{t=T+1} \rho^{t-1} f(x_{t-1}, \alpha_t(x_{t-1})) \end{aligned}$$

D. h.,

$$\left| C_\alpha^\xi - C_\alpha^{*T\xi} \right| \leq c\rho^T$$

mit einer von f, ρ abhängigen Konstanten c.

4. Schritt
Wir behaupten: $\bar{\alpha} = (\bar{\alpha}_0, \bar{\alpha}_0, \dots)$ ist optimale Strategie in Γ_Σ^ξ. Es weiche etwa Spieler 1 ab zu $\alpha^1 \in \mathfrak{S}$ („in Σ“). Dann gilt:

$$\begin{aligned} C_{(\bar{\alpha}^1, \bar{\alpha}^2)}^\xi &\geq C_{(\bar{\alpha}^1, \bar{\alpha}^2)}^{*T\xi} - c\rho^T \\ &\geq C_{(\alpha^1, \bar{\alpha}^2)}^{*T\xi} - c\rho^T \end{aligned}$$

($\bar{\alpha}$ ist optimal in Σ^{*T}, Schritt2!)

$$\geq C_{(\alpha^1, \bar{\alpha}^2)}^\xi - 2c\rho^T \qquad \text{nach Schritt 3}$$

Da T beliebig und $\rho^T \xrightarrow[T \to \infty]{} 0$ folgt

$$C_{(\bar{\alpha}^1, \bar{\alpha}^2)}^{1\xi} \geq C_{(\alpha^1, \bar{\alpha}^2)}^{1\xi}$$

– genauso für Spieler 2, also ist $\bar{\alpha}$ optimal.

5. Schritt
Natürlich ist

$$v_{\Gamma_\Sigma^\xi} = C_{\bar{\alpha}}^\xi = \lim_{t \to \infty} C_{\bar{\alpha}}^{*T\xi} = \lim_{T \to \infty} v(x) = v(x)$$

d. h., der Fixpunkt liefert Spielwert. $\square$

3 Stochastische Spiele, Unvollständige Informationen

3.1 Redeweisen der Stochastik

Ein Wahrscheinlichkeitsraum ist ein Tripel (Ω, F, m) derart, daß

1) Ω endlich oder abzählbar.

2) $\mathbf{F} = \mathfrak{P}(\Omega)$ (oder auch $\mathbf{F} \subseteq \mathfrak{P}(\Omega)$)

3) m ist eine additive Funktion auf $\mathbf{F}$, ≥ 0 und auf 1 normiert, d. h.

$$\begin{aligned} m\colon \mathbf{F} &\to [0,1] \\ m(F \cup G) &= m(F) + m(G) \qquad \text{falls } F, G \text{ disjunkt} \end{aligned}$$

Daher ist m völlig beschrieben durch $(m_\omega)_{\omega\in\Omega}$ mit $m_\omega := m(\{\omega\})$ und zwar vermöge

$$m(F) := \sum_{\omega\in F} m_\omega \quad \forall F \in \mathbf{F}$$

Interpretation: Ω Ergebnisse eines Experiments, Stichprobenraum, sample space. $F \in \mathbf{F}$ ist ein *Ereignis*. $m(F)$ ist die Wahrscheinlichkeit für das Auftreten von F.
Für abzählbare Ω ist $(m_\omega)_{\omega\in\Omega}$ eine konvergente Reihe:

$$\begin{aligned} \sum_{\omega\in\Omega} m_\omega &= 1 = m(\Omega) \\ \sum_{\omega\in F} m_\omega &= m(F) \end{aligned}$$

Ein Zufallsversuch ist eine Abbildung $f\colon \Omega \to \mathbb{R}$. Jedes m konstruiert ein Integral, Erwartungswert:

$$\int f dm = \int_\Omega f dm = \sum_{\omega\in\Omega} f(\omega) m(\{\omega\}) = \sum_\Omega f(\omega) m_\omega$$

D. h. mittlerer gewichteter Wert. Wichtig: Verhalten des Integrals unter Massentransport, „Variablentransformation".
Massentransport: Seien Ω, Ω^* Mengen und $\varphi\colon \Omega \to \Omega^*$ eine Abbildung. Sei $(\Omega, \mathbf{F}, m)$ Wahrscheinlichkeitsraum. Wir transportieren die Masse vermöge

$$\begin{aligned} m^*(\{\omega^*\}) &:= m(\{\omega \mid \varphi(\omega) = \omega^*\}) \qquad \forall \omega^* \in \Omega^* \\ &= \sum_{\substack{\omega \\ \varphi(\omega)=\omega^*}} m_\omega \\ &= m(\varphi^{-1}(\{\omega^*\})) \end{aligned}$$

Weil nun φ^{-1} mit Mengenoperationen verträglich ist ($\varphi^{-1}(S \cap T) = \varphi^{-1}(S) \cap \varphi^{-1}(T)$ usw.) haben wir eine Mengenfunktion m^* vermöge

$$m^*(S^*) = m(\varphi^{-1}(S^*)) \qquad \forall S^* \subseteq \Omega^*$$

definiert. Die Abbildung φ wirkt also auf m vermöge

$$\varphi m = m \circ \varphi^{-1} = m(\varphi^{-1}(\bullet))$$

φ transportiert Wahrscheinlichkeiten, Maße etc. kovariant. Übrigens: φ transportiert Funktionen kontravariant: Ist $f^* \colon \Omega^* \to \mathbb{R}$ eine Funktion (Zufalssvariable), so ist durch

$$f(\omega) = f^*(\varphi(\omega)) = f^* \circ \varphi(\omega)$$

eine Funktion auf Ω definiert. Die Wirkung von φ ist also durch

$$\varphi f^* = f^* \circ \varphi$$

beschrieben und ist kontravariant.

Satz 3.1.1 (Variablentransformation). *Sei* $(\Omega, \mathbf{F}, m)$ *Wahrscheinlichkeitsraum,*

$$\varphi \colon \Omega \to \Omega^* \text{ und } f^* \colon \Omega^* \to \mathbb{R} \text{ integrabel bzgl. } m^* = \varphi m.$$

Dann gilt

$$\int_\Omega (\varphi f^*) dm = \int_{\Omega^*} f^* d(\varphi m)$$

d. h.

$$\begin{aligned} \int_\Omega f^* \circ \varphi \, dm &= \int_{\Omega^*} f^* \, d(m \circ \varphi^{-1}) \\ \iff \qquad \int_\Omega f \, dm &= \int_{\Omega^*} f^* \, dm^* \end{aligned}$$

Beweis. Es ist

$$\begin{aligned} \int f \, dm &= \sum_\omega f(\omega) m_\omega = \sum_\omega f^*(\varphi(\omega)) m_\omega \\ &= \sum_{\omega^* \in \Omega^*} \sum_{\{\omega \,|\, \varphi(\omega) = \omega^*\}} f^*(\omega^*) m_\omega \\ &= \sum_{\omega^*} f^*(\omega^*) \sum_{\{\omega \,|\, \varphi(\omega) = \omega^*\}} m_\omega = \sum_{\omega^*} f^*(\omega^*) m^*_{\omega^*} \\ &= \int_{\Omega^*} f^* \, dm^*_{\omega^*} \end{aligned}$$

$\square$

Speziell sei $\Omega^* \subseteq \mathbb{R}$, $f^* = \mathrm{id}$. Dann ist

$$f = \varphi\,\mathrm{id} = \mathrm{id} \circ \varphi = \varphi$$

also gilt

$$\int_\Theta f\,dm = \int_{\Omega^*} \mathrm{id}\; d(fm)$$

$p = fm$ heißt die *Verteilung von f unter m*. Das erklärt z. B. für Normalverteilte Zufallsvariable $f\colon \Omega \to \mathbb{R}$

$$\int f\,dm = E_m f = \frac{1}{\sqrt{2\pi\sigma^2}} \int x e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx$$

3.1.1 Die bedingte Erwartung

Sei $(\Omega, \mathbf{F}, m)$ ein Wahrscheinlichkeitsraum. Sei

$$\Omega = F_1 \cup \ldots \cup F_l$$

eine Zerlegung ($F_\lambda \cap F_{\lambda'} = $) sowie

$$\mathbf{F}_0 := \{ \bigcup_{\lambda\in\Lambda_0} F_\lambda \mid \Lambda_0 \subseteq \{1, \ldots, l\}\}$$

die erzeugte Algebra,. . . Körper. Die F_λ heißen *Blöcke,* die $F \in \mathbf{F}_0$ meßbare Mengen. Auch auf $\mathbf{F}_0$ kann man additive Mengenfunktion definieren, etwa

$$\begin{aligned} m(F) &= \sum_{\lambda\in\Lambda_0} m(F_\lambda) \\ &= \sum_{\lambda\in\Lambda_0} m_\lambda \end{aligned}$$

d. h. durch Angabe von $(m_\lambda)_{\lambda=1,\ldots,l}$

Satz 3.1.2. *über bedingte Erwartung Sei $\mathbf{F}_0$ eine Algebra und $f\colon \Omega \to \mathbb{R}$ ein Zufallsvariable. Dann gibt es eine Zufallsvariable $g\colon \Omega \to \mathbb{R}$ derart, daß*

1) g ist konstant auf Blöcken (meßbar bzgl. $\mathbf{F}_0$)

2) Für jedes $F \in \mathbf{F}_0$ gilt

$$\int_F f\,dm = \int_F g\,dm$$

nämlich

$$g := \sum_{\substack{\lambda \\ m(F_\lambda)>0}} \left(\tfrac{1}{m(F_\lambda)} \int_{F_\lambda} f\,dm \right) 1_{F_\lambda}$$

g heißt bedingte Erwartung von f relativ zu $\mathbf{F}_0$, geschrieben $E(f|\mathbf{F}_0)$ mit

$$1_F(\bullet) = \begin{cases} 1 & \bullet \in F \\ 0 & \textit{sonst} \end{cases}$$

Erläuterungen: g meßbar bezüglich $\mathbf{F}_0$*:* $g(\omega) = g(\omega')$ $\omega, \omega' \in F_\lambda$*, d. h.* $g = \sum_\lambda c_\lambda 1_{F_\lambda}$ *mit passendem* $c_\lambda, \lambda = 1, \dots, l$*. Auch:* $g^{-1}(S) \in \mathbf{F}_0 \forall S \subseteq \mathbb{R}$*.*
Nicht eindeutig: $\int 1_{F_\lambda}\, dm = m(F_\lambda)$ *auf Blöcken* F_λ *mit* $m(F_\lambda) = 0$ *kann man* $E(F_\lambda) = 0$ *beliebig fortsetzen.*

Satz 3.1.3. *Die bedingte Erwartung leistet:*

1) *E ist linear:*
$$E(\alpha f + \beta g|\mathbf{F}_0) = \alpha E(f|\mathbf{F}_0) + \beta E(g|\mathbf{F}_0)$$
für Funktionen f, g *und* $\alpha, \beta \in \mathbb{R}$*.*

2) *E ist monoton.*
$$f \geq g \;\Rightarrow\; E(f|\mathbf{F}_0) \geq E(g|\mathbf{F}_0)$$

3) $E(g \cdot f|\mathbf{F}_0) = gE(f|\mathbf{F}_0)$ *mit g meßbar bzgl.* $\mathbf{F}_0$*. Insbesondere:* $E(g|\mathbf{F}_0) = g$ *für g meßbar bezgl.* $\mathbf{F}_0$ *und* $E(E(f|\mathbf{F}_0)|\mathbf{F}_0) = E(f|\mathbf{F}_0)$ *(die* $\mathbf{F}_0$*-meßbaren Funktionen bilden linearen Unterraum und* $E(\bullet|\mathbf{F}_0)$ *ist die Projektion).*

Lemma 3.1.4 (Faktorisierungssatz). *Sei* $h\colon \Omega \to \mathbb{R}$ *eine Funktion, Zufallsvariable. Sei* $\mathbf{F}_0 = \sigma(h)$ *die von h erzeugte Algebra (d. h., falls* $h = \sum_{\lambda=1}^{l} d_\lambda 1_{F_\lambda}$*, so von* $\{F_\lambda\}_{\lambda=1,\dots,l}$ *erzeugt.). Dann ist* $f\colon \Omega \to \mathbb{R}$ *meßbar bzgl.* $\sigma(h)$ *genau dann, wenn es eine Funktion* $k\colon \mathbb{R} \to \mathbb{R}$ *gibt mit*
$$f = k \circ h.$$
$\sigma(h) =$ *von* $F_1, \dots, F_l$ *erzeugt,* $F_\lambda = \{\omega \mid h(\omega) = d_\lambda\}$

Beweis. Ist $f = \sum e_\lambda 1_{F_\lambda}$, so setze $k(d_\lambda) = e_\lambda \;\forall \lambda \in \{1, \dots, l\}$, $k(x)$ beliebig sonst. Dann ist

$$\begin{aligned} f(\omega) &= k(h(\omega)) \\ &= k(d_\lambda) = e_\lambda \quad \forall \omega \in F_\lambda \end{aligned}$$

□

Da $E(f|\sigma(h))$ meßbar ist bezgl. $\sigma(h)$ existiert $k\colon \mathbb{R} \to \mathbb{R}$ mit

$$E(f|\sigma(h)) = k \circ h. \tag{H}$$

Darum führt man Notation ein:

$$\begin{aligned} E(f|h) &:= E(f|\sigma(h)) \qquad \text{ist Fkt., Zufallsvariable} \\ E(f|h = x) &:= k(x) \in \mathbb{R} \end{aligned}$$

falls (H) gilt. Man sieht leicht:

$$E(f|h = d_\lambda) = \frac{1}{m(F_\lambda)} \int_{F_\lambda} f\, dm$$

falls $F_\lambda = \{\omega \mid h(\omega) = d_\lambda\}$.

Wir haben Versionen der bedingten Erwartung:

$$\begin{aligned} E(f|\mathbf{F}_0) &\quad \text{Zufallsverteilung} \\ E(f|h) &\quad \text{Zufallsverteilung} \\ E(f|h=x) &\quad \in \mathbb{R} \end{aligned}$$

weitere Versionen sind:

$$\begin{aligned} E(f|T) &:= E(f|1_T = 1) \qquad (E \in \mathbf{F}) \\ m(S|\dots) &:= E(1_S|\dots) \qquad \text{bedingte Wahrscheinlichkeit} \end{aligned}$$

z. B.:

$$\begin{aligned} m(S|\mathbf{F}_0) &:= \sum_{m(F_\lambda)>0} \left(\frac{1}{m(F_\lambda)} \int_{F_\lambda} 1_S \, dm \right) 1_{F_\lambda} & \int_{F_\lambda} 1_S \, dm &= \sum_{\omega \in F_\lambda} 1_S(\omega) m_\omega \\ &= \sum_{m(F_\lambda)>0} \frac{m(S \cap F_\lambda)}{m(F_\lambda)} 1_{F_\lambda} & &= \sum_{S \cap F_\lambda} m_\omega \\ & & & m(F_\lambda \cap S) \end{aligned}$$

und

$$\begin{aligned} m(S|T) &= E(1_S|1_T = 1) \qquad (\sigma(1_T) = \{, T, \complement T, \Omega\}) \\ &= \tfrac{\int_T 1_S \, ds}{m(T)} \\ &= \frac{m(S \cap T)}{m(T)} \qquad \text{naive bedingte Wahrscheinlichkeit} \end{aligned}$$

3.1.2 Spezielle Maße auf Produkträumen

Seien $(\Omega_1, \mathbf{F}_1, m^1)$, $(\Omega_2, \mathbf{F}_2, m^2)$ Wahrscheinlichkeitsräume. Die Produktwahrscheinlichkeit

$$m = m^1 \otimes m^2$$

auf $(\Omega_1 \times \Omega_2, \mathbf{F}_{12})$ ist definiert durch

$$m_{(\omega_1,\omega_2)} = m_{\omega_1} m_{\omega_2} \qquad (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$$

m ist Wahrscheinlichkeit!
Allgemeiner: Sei

$$\underline{\mathsf{X}} = \underline{\mathsf{X}}_0 \times \cdots \times \underline{\mathsf{X}}_T$$

sowie für jedes t stets m^t eine Wahrscheinlichkeit auf $\underline{\mathsf{X}}_t$. Das Produktmaß ist

$$m = m^0 \otimes \cdots \otimes m^T$$

gegeben durch

$$m_x = m^0_{x_0} \cdot \cdots \cdot m^T_{x_T} \qquad x \in \underline{\mathsf{X}}, x = (x_0, \dots, x_T)$$

Genauer $m(\{x\}) = m_x$, $m(F) = \sum_{x \in F} m_x$. ($m$ ist Wahrscheinlichkeitsverteilung!)

Weitere Konstruktionen: Markoff-Maße. Schreibweise: ein *stochastischer Kern*, geschrieben

$$P|\Omega_1 \Rightarrow \Omega_2$$

ist eine Abb.

$$P\colon \Omega_1 \times \mathfrak{P}\Omega_2 \to [0,1]$$

so daß für jedes ω_1 stets $P(\omega, \bullet)$ eine Wahrscheinlichkeitsverteilung ist. (Natürlich ist für Ω_1, Ω_2 endl.

$$P(\omega_1, \omega_2) = P(\omega_1, \{\omega_2\})$$

mit ≥ 0, $\sum_{\omega_2} P(\omega_1, \omega_2) = 1$). $P(\bullet, \bullet)$ ist stochastische Matrix.
P ordnet jedem ω_1 eine Wahrscheinlichkeitsverteilung $P(\omega_1, \bullet)$ zu.
Sei wieder

$$\underline{\overline{\mathsf{X}}} = \underline{\overline{\mathsf{X}}}_0 \times \cdots \times \underline{\overline{\mathsf{X}}}_T$$

sowie μ^0 eine Wahrscheinlichkeitsverteilung auf $\underline{\overline{\mathsf{X}}}_0$. Ferner sei für jedes $t \in T$ ein stoch. Kern $P_t|\underline{\overline{\mathsf{X}}}_{t-1} \Rightarrow \underline{\overline{\mathsf{X}}}_t$ gegeben. Das *Markoffsche Maß* $m = m_{\mu^0}^{P_\bullet}$ ist gegeben durch

$$m_x := \mu^0_{x_0} P_1(x_0, x_1) \ldots P_T(x_{T-1}, x_T) \qquad \forall x \in \underline{\overline{\mathsf{X}}}$$

1) $m_{\mu^0}^{P_\bullet}$ ist Wahrscheinlichkeitsverteilung!

2) Produktmaße sind spezielle Markoff-Maße

Sei $(\Omega, \mathbf{F}, m)$ ein Wahrscheinlichkeitsraum. Ein Prozeß ist eine Abbildung $x\colon \Omega \to \underline{\overline{\mathsf{X}}}$. xm ist die Verteilung von x unter m. ($xm = m \circ x^{-1}$)
x heißt *Markoffscher Prozeß,* falls $xm = m_{\mu^0}^{P_\bullet}$ ist. Es reicht i. a., die Verteilung zu kennen!
Z. B.: Welches ist die Wahrscheinlichkeit, daß der Prozeß $\bar{x}_t \in \underline{\overline{\mathsf{X}}}_t$ trifft? D. h., was ist

$$m(\{\omega | x_t(\omega) = \bar{x}_t\})?$$

d. h., mit $S = \{x | x = (x_0, \ldots, x_T), x_t = \bar{x}_t\}$ ist gesucht

$$\begin{aligned} m(\{\omega | x(\omega) \in S\}) &= m \circ x^{-1}(S) = (Xm)(S) \\ &= m_{\mu^0}^{P_\bullet}(S) \end{aligned}$$

Dies ist zu berechnen wie folgt

$$m_{\mu^0}^{P_\bullet} = \sum_{\substack{x \in \underline{\overline{\mathsf{X}}} \\ x_t = \bar{x}_t}} \mu^0_{x_0} P_1(x_0, x_1) \cdots P_T(x_{T-1}, x_T)$$

Summiere die letzten Terme zu 1

$$\begin{aligned} &= \sum_{x_0 \ldots x_{t-1}} \mu^0_{x_0} P_1(x_0, x_1) \cdots P_t(x_{t-1}, \bar{x}_t) \\ &= (\mu^0 P_1 \cdots P_T)_{\bar{x}_t} \end{aligned}$$

D. h., die Verteilung von x_t ist $x_t m = \mu^t = \mu^0 P_1 \cdots P_t$.
Markoffscher Prozeß:

- Vorgegeben $\mu^0, P_1, \ldots, P_T$
- Konstruktion von $m^P_{\mu^0}$ auf $\overline{\underline{X}}$
- x Markoff-Prozeß. Verteilung $xm = x^{P_\bullet}_{\mu^0}$ ist Markoffsch

$$x\colon \Omega \to \overline{\underline{X}}$$

- Berechnung interessierender Wahrscheinlichkeiten:

$$\begin{aligned} m(\{x_t = \bar{x}_t\}) &= (\mu^0 P_1 \cdots P_t)_{\bar{x}_t} \\ &=: \mu^t_{\bar{x}_t} \end{aligned}$$

  μ^t ist Verteilung von x^t!

Ähnlich:

$$\begin{aligned} m(\{\omega | x_t\bullet = \bar{x}_t, x_{t-1}\bullet = \bar{x}_{t-1}\}) &= m(\{x_t = \bar{x}_t, x_{t-1} = \bar{x}_{t-1}\}) \\ &= (\mu^0 P_1 \cdots P_{t-1})_{\bar{x}_{t-1}} P_t(\bar{x}_{t-1}, \bar{x}_t) \\ &= \mu^{t-1}_{\bar{x}_{t-1}} P_t(\bar{x}_{t-1}, \bar{x}_t) \end{aligned}$$

D. h., $\mu^{t-1}_\bullet P_t(\bullet, x)$ ist Verteilung der Funktion

$$(x_{t-1}, x_t)\colon \Omega \to \overline{\underline{X}}_{t-1} \times \overline{\underline{X}}_t$$

(gemeinsame Verteilung von x_{t-1}, x_t.)

$$\begin{aligned} m(\{x_t = \bar{x}_t\} | \{x_{t-1} = \bar{x}_{t-1}\}) &= \frac{m(\{x_t = \bar{x}_t, x_{t-1} = \bar{x}_{t-1}\})}{m(\{x_{t-1} = \bar{x}_{t-1}\})} \\ &= \frac{\mu^{t-1}_{\bar{x}_{t.1}} P_t(\bar{x}_{t-1}, \bar{x}_t)}{\mu^{t-1}_{\bar{x}_{t-1}}} \\ &= P_t(\bar{x}_{t-1}, \bar{x}_t) \end{aligned}$$

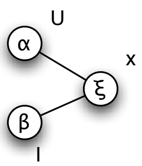
D. h. Vorgabe von μ^0 und $P_1, \ldots, P_T$ bedeutet Vorgabe von Anfangsverteilung und bedingten Wahrscheinlichkeiten des Übergangs. 8.Juli fehlt

3.2 Kausale Netzwerke

13. Juli

$(\mathfrak{X}, \prec, Z, P_\bullet)$ (Graph, Zustände) $\overline{\underline{Z}}_\xi = \{0, \ldots, a_\xi\}$, $\overline{\underline{Z}} = (\overline{\underline{Z}}_\xi)_{\xi \in \mathfrak{X}}$
$\overline{\underline{Z}}_v = \prod_{\eta \prec \xi} (\overline{\underline{Z}}_\eta) = \overline{\underline{Z}}_{v(\xi)}$, $P_\xi | \overline{\underline{Z}}_v \Rightarrow \overline{\underline{Z}}_\xi$: Jedem $(z_\xi)_{\xi \in v}$ ist eine Wahrscheinlichkeitsverteilung $P_\xi((z_\xi)_{\xi \in v}, \bullet)$ zugeordnet.

Beispiel 3.2.1. Verursacher U, Inhibitor I (Störung), x Resultat, alle $z_\xi = \{0, 1\}$.



1) Vorstellungen: $(x_1, U), (x_2, I)$ stoch. unabhängig, $x = x_1 x_2$, d. h.

$$P(x = 1 \mid U, I) = P(x_1 = 1 \mid U) P(x_2 = 1 \mid I)$$

also

$$\begin{aligned} P(x = 1 \mid U = 1, I = 1) &= 1 \cdot q = q \\ P(x = 1 \mid U = 1, I = 0) &= 1 \cdot 1 = 1 \end{aligned}$$

2) Dem entspricht Ansatz für Markoffsche Kerne. Wir definieren

$$P_{\alpha\xi} | \underline{\overline{Z}}_\alpha \Rightarrow \underline{\overline{Z}}_\xi, \quad P_{\beta\xi} | \underline{\overline{Z}}_\beta \Rightarrow \underline{\overline{Z}}_\xi$$

und setzen

$$\begin{aligned} P_{\alpha\xi}(1, \bullet) &= (1, 0) \\ P_{\beta\xi}(1, \bullet) &= (q, 1 - q) \end{aligned}$$

Der Markoffsche Kern

$$P_\xi | \underline{\overline{Z}}_v \Rightarrow \underline{\overline{Z}}_\xi$$

wird nun zusammengesetzt:

$$P_\xi(z_\bullet, 1) := P_{\alpha\xi}(z_\alpha, 1) P_{\beta\xi}(z_\beta, 1)$$

3) Gegeben $\mu^\alpha_\bullet$ und $\mu^\beta_\bullet$ (Anfangsverteilungen), kann man nun das Markoffmaß $m^{P_\bullet}_{\mu_\bullet}$ konstruieren wie im Konstruktionssatz angegeben. Danach ist die Verteilung eines Prozesses $Z \colon \Omega \to \underline{\overline{Z}}$ bekannt (nämlich $m^{P_\bullet}_{\mu_\bullet}$) – es ist also sinnvoll,

$$P(x = 1 \mid U = 1, I = 0)$$

zu schreiben.

4) Typisch möchte man nun Wahrscheinlichkeiten interessierender Art berechnen. Z. B.: Was ist die Wahrscheinlichkeit einer Fehlfunktion, wenn der ξ-Schalter offen ist? (I.A. „Bayes'scher Satz“)

Gefragt ist

$$\begin{aligned}
P(I=1 \mid x=1) &\overset{!}{=} P(U=1, I=1 \mid x=1) \\
&= \frac{P(U=1, I=1, x=1)}{P(x=1)} \\
&= \frac{P(U=1, I=1, x=1) P(U=1, I=1)}{P(U=1, I=1) P(x=1)} \\
&= P(x=1 \mid U=1, I=1) \frac{P(U=1, I=1)}{P(x=1)} \\
&= \frac{q\mu_1^\beta \mu_1^\alpha}{P(x=1)} \\
&= \frac{q\mu_1^\alpha \mu_1^\beta}{\sum_{z_\alpha z_\beta} P(x=1 \mid (I,U)=(z_\alpha, z_\beta)) \cdot P((I,U)=z_\alpha z_\beta)} \\
&= \frac{q\mu_1^\beta \mu_1^\alpha}{q\mu_1^\beta \mu_1^\alpha + 1 \cdot \mu_0^\beta \mu_1^\alpha + 0} \\
&= \frac{q\mu_1^\beta}{q\mu_1^\beta + \mu_0^\beta}
\end{aligned}$$

Angeregt durch das Beispiel:

Definition 3.2.2. *Sei* $\Sigma = (\mathfrak{X}, \prec, Z, P_\bullet)$ *ein kausales Netzwerk. Sei* $\underline{\overline{Z}}_\xi = \{0,1\} \forall \xi \in \mathfrak{X}$. *Man sagt* $\xi \notin \mathfrak{R} = \{\xi \mid v(\xi) = \emptyset\}$ *konstituiere* konjunktive Interaktion *(AND-gate), falls es Kerne*

$$P_{\eta\xi} | \underline{\overline{Z}}_\eta \Rightarrow \underline{\overline{Z}}_\xi \quad \forall \eta \in v(\xi)$$

gibt, so daß

$$P_\xi(z_v, 1) = \prod_{\eta \in v(\xi)} P_{\eta\xi}(z_\eta, 1)$$

gilt. ξ *konstituiert disjunktive Interaktion, falls gilt*

$$P_\xi(z_v, 1) = 1 - \prod_{\eta \in v(\xi)} P_{\eta\xi}(1 - z_\eta, 1)$$

(OR-gate)

Lemma 3.2.3. *Angenommen, es sei bei einem OR-gate*

$$P_{\eta\xi}(0,1) = 0, \quad P_{\eta\xi}(1,1) = 1$$

Dann ist „ξ offen“ genau dann, wenn wenigstens ein Vorgänger offen ist. Denn

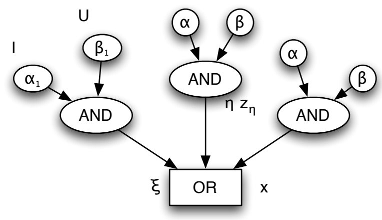
Beweis.

$$\begin{aligned}
P(z_\xi = 1 \mid z_v = z_v) &= 1 - \prod_v P_{\eta\xi}(1 - z_\eta, 1) \\
&= 1 - \prod_{\eta \in v, z_\eta = 1} P_{\eta\xi}(0,1) \prod_{\eta \in v, z_\eta = 0} P_{\eta\xi}(1,1) \\
&= \begin{cases} 1 & \text{falls wenigstens ein } z_\eta = 1 \\ 0 & \text{sonst} \end{cases}
\end{aligned}$$

□

Man kann nun zusammensetzen.

Beispiel 3.2.4 *(Das gestörte OR-gate (noisy or-gate)).* Jedes AND-gate ist wie Ursache/Inhibitor konstruiert.



Wir berechnen

$$P(x=1 \mid U=u, I=v) = \sum_{z\in \underline{\mathsf{Z}}_v} P(x=1 \mid Z=z)P(Z=z \mid U=u, I=v)$$

Wir beginnen mit

$$\begin{aligned} P(x=1 \mid Z=z) &= 1 - \prod_{\eta\in v} P(x=1 \mid Z_\eta \neq z_\eta) \\ &= 1 - \prod_{\eta\in v, z_\eta=1} P(x=1 \mid Z_\eta=0) \prod_{\eta\in v, z_\eta=0} P(x=1 \mid Z_\eta=1) \\ &= \begin{cases} 1 & \text{mind ein } z_\eta=1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Also fährt man fort

$$\begin{aligned} P(x=1 \mid U=u, I=v) &= \sum_{\substack{z\in \underline{\mathsf{Z}}_v \\ \text{mind ein } z_\eta=1}} 1\cdot P(Z=z \mid U=u, I=v) \\ &= 1 - P(Z=0 \mid U=u, I=v) \\ &= 1 - \prod_{\eta\in v} P(Z_\eta=0 \mid U_\eta=u_\eta, I_\eta=v_\eta) \\ &= 1 - \prod_{\eta\in v, u_\eta=0} 1 \prod_{\substack{\eta\in v\\ u_\eta=1\\ v_\eta=0}} 0 \prod_{\substack{\eta\in v\\ u_\eta=1\\ v_\eta=1}} q_\eta \\ &= \begin{cases} 1 & \text{falls } \exists\eta : u_\eta=1, v_\eta=0 \\ 1-\prod_{\substack{\eta\in v\\ u_\eta=1\\ v_\eta=1}} q_\eta & \text{sonst} \end{cases} \end{aligned}$$

3.3 Prozesse auf Bäumen

15. Juli 04

Kausale Netzwerke sollen spezialisiert werden zu Prozessen auf einem Spielbaum (für die stoch. dyn. Optimierung). wir versuchen also „Brückenschlag“! Vorstellung von „Schaltern“ soll Begriff von zufälligen Pfaden liefern. Stets ist $(\mathfrak{X}, \prec)$ ein Baum.

Definition 3.3.1. *Sei* $(\mathfrak{X}, \prec, Z)$ *ein Netzwerk, jedes* $\underline{\overline{Z}}_\xi = \{0,1\}$. *Sei* $(\Omega, \mathbf{F}, P)$ *ein Wahrscheinlichkeitsraum. Eine Zufallsverteilung* $Z\colon \Omega \to \underline{\overline{Z}}$ *heißt* Knotenbelegung, *falls gilt*

1) $P(Z_\eta = 1 \mid Z_\xi = 1) = 1 \qquad \forall \eta \prec \xi$

2) $\sum_{\eta \in N(\xi)} P(Z_\eta = 1 \mid Z_\xi = 1) = 1 \qquad \forall \xi \in \mathfrak{X} - \partial\mathfrak{X} = \mathring{\mathfrak{X}}$

(Gegeben ein Knoten ist „an", muß der Vorgänger „an" sein und mit Wahrscheinlichkeit 1 ein Nachfolger angeschaltet werden.)

Lemma 3.3.2. *Sei* $x = (x_0, \dots, x_T)$ *ein Pfad mit* $x_0 = \xi_0$ *(Wurzel),* $x_T \in \partial\mathfrak{X}$. *Ist* Z *eine Knotenbelegung, so gilt*

1) $P(Z_{x_T} = 1) = P(Z_{x_0} = 1, \dots, Z_{x_T} = 1)$

2) $P(Z_{x_T} = 1 \mid Z_{x_{T-1}} = 1) = P(Z_{x_T} = 1 \mid Z_{x_{T-1}} = 1, \dots, Z_{x_0} = 1)$

D. h., der letzte Knoten ist „offen", falls alle unterwegs offen sind – und Z *verhält sich Markoffsch.*

Beweis. 1) Allgemein gilt für Mengenfunktionen m (additiv):

$$m(A \cap B) = m(A) \Rightarrow m(A \cap B \cap F) = m(A \cap F)$$

wegen

$$m(A \cap F) = m(A \cap B \cap F) + \underbrace{m((A - B) \cap F)}_{=0}$$

2) Das benutzen wir: zunächst ist

$$1 = P(Z_{x_{t-1}} = 1 \mid Z_{x_t} = 1) = \frac{P(Z_{x_{t-1}} = 1, Z_{x_t} = 1)}{P(Z_{x_t} = 1)}$$

$$\begin{aligned} \text{Also} \quad & P(Z_{x_t} = 1) = P(Z_{x_{t-1}} = 1, Z_{x_t} = 1) \\ \text{und} \quad & P(Z_{x_{t-1}} = 1) = P(Z_{x_{t-1}} = 1, Z_{x_{t-2}} = 1) \end{aligned}$$

Daher (nach eben) schneide mit $\{Z_{x_t} = 1\}$

$$\begin{aligned} P(Z_{x_t} = 1, Z_{x_{t-1}} = 1) &= P(Z_{x_t} = 1, Z_{x_{t-1}} = 1, Z_{x_{t-2}} = 1) \\ &= P(Z_{x_t} = 1) \end{aligned}$$

usw. – dies zeigt die erste Behauptung.
Zu Punkt 2 (Markoffsch):
Es ist

$$P(Z_{x_T} = 1 \mid Z_{x_{T-1}} = 1) = \frac{P(Z_{x_T} = 1, Z_{x_{T-1}} = 1)}{P(Z_{x_{T-1}} = 1)}$$

nach 2. Schritt

$$\begin{aligned} &= \frac{P(Z_{x_T} = 1, \dots, Z_{x_0} = 1)}{P(Z_{x_{T-1}} = 1, \dots, Z_{x_0} = 1)} \\ &= P(Z_{x_T} = 1 \mid Z_{x_{T-1}} = 1, \dots, Z_{x_0} = 1) \end{aligned}$$

□

Satz 3.3.3. *Sei Z eine Knotenbelegung. Es sei $P(Z_{\xi_0} = 1) = 1$. Dann ist*

$$\sum_{x \in \underline{\overline{\mathsf{X}}}} P(Z_{x_0} = 1, \ldots, Z_{x_T} = 1) = 1$$

d. h. mit Wahrscheinlichkeit 1 geht's irgendo entlang, wenn bei ξ_0 „an“ war.

Beweis. durch Induktion. Nach Annahme

$$\sum_{x_0} P(Z_{x_0} = 1) = P(Z_{\xi_0}) = 1$$

Für Länge 2:

$$\begin{aligned}\sum_{\{x_1 \mid x_0 \prec x_1\}} P(Z_{x_0} = 1, Z_{x_1} = 1) &= \sum_{\ldots} \frac{P(Z_{x_0} = 1, Z_{x_1} = 1) P(Z_{x_0} = 1)}{P(Z_{x_0} = 1)} \\ &= \underbrace{\sum_{\ldots} P(Z_{x_1} = 1 \mid Z_{x_0} = 1)}_{=1 \text{ n. Def.}} \underbrace{P(Z_{x_0} = 1)}_{=1 \text{ n. Vor.}}\end{aligned}$$

usw. $\square$

Definition 3.3.4. *Sei Z eine Knotenbelegung. $\omega \in \Omega$ erzeugt einen Pfad $x \in \underline{\overline{\mathsf{X}}}$, falls*

$$\{\xi \mid Z_\xi(\omega) = 1\} = \{x_0, \ldots, x_T\}$$

Satz 3.3.5. *Sei Z eine Knotenbelegung. Es gelte*

1) $P(Z_{\xi_0} = 1) = 1$

2) *für alle $\xi \in \mathring{\mathfrak{X}}$ sei*

$$P(Z_\eta = 1, Z_{\eta'} = 1) = 0 \qquad \forall \eta \neq \eta' \in N(\xi)$$

Dann existiert $\bar{\Omega} \subseteq \Omega$ mit

1) $P(\bar{\Omega}) = 1$

2) *Jedes $\omega \in \bar{\Omega}$ erzeugt einen Pfad.*

D. h., sind Verzweigungen verboten, so erzeugt die Belegung mit Wahrscheinlichkeit 1 Pfade.

Beweis. Für jedes $x \in \underline{\overline{\mathsf{X}}}$ sei

$$\begin{aligned}\Omega_x &:= \{\omega \mid \{\xi \mid Z_\xi(\omega) = 1\} = \{x_0, \ldots, x_T\}\} \\ &= \{\omega \mid \omega \text{ erzeugt } x\}\end{aligned}$$

Nach einem früheren Satz wird irgendein Pfad mit Sicherheit erzeugt. Also ist

$$P\Big(\bigcup_{x \in \underline{\overline{\mathsf{X}}}} \Omega_x\Big) = 1$$

Wegen der zweiten Voraussetzung ist $P(\Omega_x \cap \Omega_{x'}) = 0$ für $x \neq x'$. Setzt man daher

$$\bar{\Omega}_x := \Omega_x - \bigcup_{x' \neq x} \Omega_{x'}$$

so ist $P(\bar{\Omega}_x) = P(\Omega_x)$ und $P(\bigcup \bar{\Omega}_x) = 1$, $P(\bar{\Omega}_x \cap \bar{\Omega}_{x'}) = 0$. Sogar $\bar{\Omega}_x \cap \bar{\Omega}_{x'} = \emptyset$ für $x \neq x'$. Daher leistet $\bar{\Omega} := \bigcup_{x \in \underline{\mathrm{X}}} \bar{\Omega}_x$ das verlangte. Es gibt offnebar zu jedem $\omega \in \bar{\Omega}$ genau ein x mit „ω erzeugt x“, d. h.,

$$\{\xi \mid Z_\xi(\omega) = 1\} = \{x_0, \ldots, x_T\}$$

□

Definition 3.3.6. *Ein* Prozeß *ist eine Abb.*

$$X : \Omega \to \underline{\mathrm{X}}$$

Ein Fluß *ist eine Knotenbelegung Z, die die Voraussetzungen des Satzes erfüllt. Z induziert X, wenn $\{\xi \mid Z_\xi(\omega) = 1\} = \{X_0(\omega), \ldots, X_T(\omega)\} \quad \forall \omega \in \Omega$ gilt.*

Satz 3.3.7. *Jeder Fluß induziert eindeutig einen Prozeß. Jeder Prozeß wird von genau einem Fluß induziert.*

Satz 3.3.8 (Über die Konstruktion einer Verteilung). *Sei $(\mathfrak{X}, \prec)$ ein Baum. Für jedes $\xi \in \mathfrak{X}$ sei eine Wahrscheinlichkeitsverteilung $\mu^\xi_\bullet$ auf $N(\xi)$ definiert. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathbf{F}, P)$ und ein Fluß Z mit*

1) $P(Z_\eta = 1 \mid Z_\xi = 1) = \mu^\xi_\eta \quad \forall \xi \prec \eta$

2) Der induzierte Prozeß X leistet

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_T = x_T) = \mu^{x_0}_{x_1} \mu^{x_1}_{x_2} \ldots \mu^{x_{T-1}}_{x_T} \quad \forall x \in \underline{\mathrm{X}} \qquad (*)$$

Beweis. Setze $\Omega = \underline{\mathrm{X}}$ und definiere eine Verteilung m durch $(*)$, $X : \underline{\mathrm{X}} \to \underline{\mathrm{X}}$ ist die Identität.

□

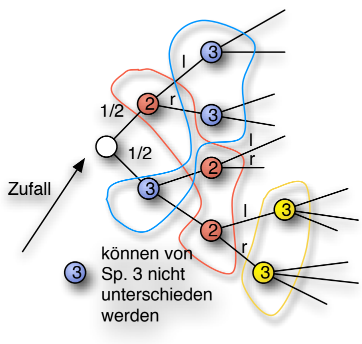
3.4 Imperfekte / Unvollständige Information

Wir wollen Spiele beschreiben mit

1) Zufallszügen

2) Imperfekte Info: verschiedene Information der Spieler über den vorliegenden Zustand

Anschaulich:



Zunächst ist die extensive Form aufzustellen:

Definition 3.4.1. *Ein* Spielbaum mit imperfekter Information *ist ein 8-Tupel*

$$\Sigma = (\mathfrak{X}, \prec, \iota, \mu^\bullet, \kappa, \sigma, f, u)$$

mit folgenden Einzelheiten:

- $(\mathfrak{X}, \prec, \iota, f, u)$ *ist ein Spielbaum wie in Abschnitt 1.2. Insbesondere ist*

  $$\iota\colon \mathfrak{X} - \partial\mathfrak{X} \to \{0, 1, \dots, n\} = \{0\} \cup I$$

  die Spieler-Knoten-Zuordnung, wobei „0" der Zufall ist.

- $\mu^\bullet = (\mu^\xi)_{\xi\in\mathfrak{X}^0=\iota^{-1}(\{0\})}$ *ist eine Familie von Wahrscheinlichkeitsverteilungen – jedes* μ^ξ *ist eine Wahrscheinlichkeitsverteilung* $(\mu^\xi_\eta)_{\eta\in N(\xi)}$ *auf* $N(\xi)$*. (Kommt eine Partei bei* ξ *an, so wird ein Zufallsmechanismus betätigt, der den Nachfolger liefert)*

- *Ferner ist* $\kappa = (\kappa^i)_{i\in I}$ *eine Familie von Abbildungen*

  $$\kappa^i\colon \mathfrak{X}^i \to \mathfrak{Q}^i$$

  dabei ist

  $$\mathfrak{X}^i := \iota^{-i}(\{i\})$$

  und $\mathfrak{Q}^i$ *eine endliche Menge (Spieler is Signale) (κi ist „Signalabbildung", an* ξ *erfährt Spieler* i *das Signal* $\kappa^i(\xi)$*) Beachte, daß*

  $$\mathbf{Q}^i := \{(\kappa^i)^{-1}(q) \mid q \in \mathfrak{Q}^i\}$$

  eine Zerlegung von $\mathfrak{X}^i$ *konstituiert. Jedes* $Q \in \mathbf{Q}^i$ *heißt Info-Menge von* i*.*

- σ *klärt, wie Spieler* i *über den Fortgang entscheidet: Für jedes* $q \in \mathfrak{Q}^i$ *und jedes* $\xi \in (\kappa^i)^{-1}(\{q\})$ *gibt es eine endliche Menge* S^q *(choices, Auswahl, Alphabet) und eine bijektive Abbildung*

  $$\sigma_\xi\colon N(\xi) \to S^q,$$

  es ist $\sigma = (\sigma_\xi)_{\xi\in\mathfrak{X}-\partial\mathfrak{X}}$*, d. h., observiert Spieler* i *das Signal* q*, so kann er* $s \in S^q$ *qählen und der nächste Knoten ist* $\sigma_\xi^{-1}(s) \in N(\xi)$*)*

Bemerkung 3.4.2. Jedes σ_ξ bildet $N(\xi)$ bijektiv auf S_q ab, also müssen alle $\xi \in Q = (\kappa^i)^{-1}(q)$ notwendig die gleiche Anzahl von Nachfolgern haben. (Wegen endl. Mengen; natürlich werden auch die Nachfolger durch σ_ξ identifiziert: $C_l := \{\sigma_\xi^{-1}(l) \mid \xi \in Q\}$.)

Nun: strategisches Verhalten:

Definition 3.4.3. *Eine* reine Strategie *für Spieler* i *ist eine Abbildung*

$$\alpha^i\colon \mathfrak{Q}^i \to \bigcup_{q\in\mathfrak{Q}^i} S^q$$

mit $\alpha^i(q) \in S^q \quad \forall q \in \mathfrak{Q}^i$*.* $\mathfrak{S}^i := \{\alpha^i \mid \alpha^i \text{ ist reine Str.}\}$*,* $\mathfrak{S} = \mathfrak{S}^i \times \cdots \times \mathfrak{S}^n$ *wir wie üblich benützt.*

Definition 3.4.4. *Eine* Verhaltensstrategie *(behavioral strategy) für Spieler i ist eine Familie*

$$A^i = (A^i(q, \bullet))_{q \in \mathfrak{Q}^i},$$

so daß für jedes q stets $A^i(q, \bullet)$ eine Wahrscheinlichkeitsverteilung auf S^q ist. D. h., Spieler i „randomisiert“ seine Entscheidung bei jedem $q \in \mathfrak{Q}^i$. $\mathfrak{A}^i = \{$ Verhaltensstrategien$\}$

Definition 3.4.5. *Eine gemischte Strategie für Spieler i ist eine Wahrscheinlichkeitsverteilung über $\mathfrak{S}^i$. M^i bezeichnet „mixed strategies“, $\mathfrak{M}^i$ die Menge der gemischten Strategien.*

Definition 3.4.6. *Sei $A \in \mathfrak{A} = \mathfrak{A}^1 \times \ldots \times \mathfrak{A}^n$ ein n-Tupel von Verhaltensstrategie. Dann ist die m_μ^A eine Wahrscheinlichkeitsverteilung auf $\underline{\overline{\mathsf{X}}}$ wie folgt gegeben:*

1) *Setze $q_\eta^\xi := A^{\iota(\xi)}(\kappa^{\iota(\xi)}(\xi), \sigma_\xi(\eta)) \quad \forall \xi \in \mathfrak{X} - \partial\mathfrak{X}, \eta \in N(\xi)$ (insbesondere: $q_\eta^\xi = \mu_\eta^\xi$ für $\xi \in \mathfrak{X}^0$)*

2) *Definiere eine Verteilung einer Knotenbelegeung Z durch*

$$P(Z_\eta = 1 \mid Z_\xi = 1) = q_\eta^\xi$$

3) *Äquivalent sei $m_\mu^A(\{x\}) = q_{x_1}^{x_0} \times q_{x_2}^{x_1} \cdots q_{x_T}^{x_{T-1}}$ und $X \colon \Omega \to \underline{\overline{\mathsf{X}}}$ der zu Z gehörige Prozeß mit Verteilung m_μ^A.*

Statt $\alpha \to x^\alpha$ nun $A \to m_\mu^A$ auf $\underline{\overline{\mathsf{X}}}$. Z. B. ist m_μ^A die Verteilung eines Prozesses x, der durch eine fließende Knotenbelegung Z erzeugt ist mit

$$P(Z_\eta = 1 \mid Z_\xi = 1) = \begin{cases} A^i(x^i(\xi), \sigma_\xi(\eta)) & \xi \in \mathfrak{X}^i \\ \mu_\eta^\xi & \xi \in \mathfrak{X}^0, \eta \in N(\xi) \end{cases}$$

Definition 3.4.7. *Die Auszahlung an Spieler i bei A ist*

$$C_A^{i\mu} = \int_{\underline{\overline{\mathsf{X}}}} \sum_{t=1}^{\tau_0} f^i(x_{t-1}, x_t) + u^i(x_{\tau_0})\, dm_\mu^A(x)$$

mit $\tau_0 \colon \underline{\overline{\mathsf{X}}} \to \mathbb{N}, \tau_0(x_1, \ldots, x_T) = T$

Natürlich ist für jeden Prozeß X mit Verteilung m_μ^A stets

$$C_A^{i\mu} = \int_\Omega \sum_{t=1}^{\tau} f^i(X_{t-1}, X_t) + u^i(X_\tau)\, dP$$

(mit $\tau = \tau_0 \circ x$) wegen der Variablentransformation.

Bemerkung 3.4.8. Eine reine Strategie ist eine spezielle Verhaltensstrategie. Ist δ_ω die auf ω konzentrierte Wahrscheinlichkeitsverteilung ($\delta_\omega(F) = \begin{cases} 1 & \omega \in F \\ 0 & \text{sonst} \end{cases}$), so liefert jede reine Strategie α^i eine Verhaltensstrategie

$$A^i(q, \bullet) = \delta_{\alpha^i(q)}(\bullet)$$

(Ebenso: eine reine Strategie ist eine spezielle gemischte)

Daher ist für jedes reine α die erzeugte Verteilung m_μ^α bereits durch die Diskussion über Verhaltensstrategie erklärt. Insbesondere ist für Spiele ohne Zufallszüge eine Abbildung $\alpha \to x^\alpha$ wohldefiniert und

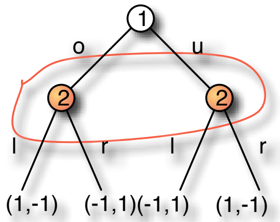
$$m_\mu^\alpha = \delta_{x^\alpha}(\bullet)$$

Definition 3.4.9. *Die Normalform in Verhaltensstrategien ist ist*

$$\Gamma_{\Sigma^\mu} = (\mathfrak{A}^1, \dots, \mathfrak{A}^n, C_A^{1\mu}, \dots, C_A^{n\mu})$$

in reinen Strategien

$$\tilde{\Gamma}_{\Sigma^\mu} = (\mathfrak{S}^1, \dots, \mathfrak{S}^n, C_\bullet^{1\mu}, \dots, C_\bullet^{n\mu})$$



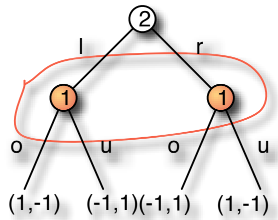
Beispiel 3.4.10.
Für perfekte Info in reinen Strategien

$$C_\bullet^1 = \begin{array}{ccccc} & ll & lr & rl & rr \\ o & 1 & \dots & & \\ u & \dots & \dots & & \end{array}$$

jetzt $1, 1$ für Spieler 2

$$C_\bullet^1 = \begin{array}{ccc} & l & r \\ o & 1 & -1 \\ u & -1 & 1 \end{array}$$

$$C_\bullet^2 = \begin{array}{ccc} & l & r \\ o & -1 & 1 \\ u & 1 & -1 \end{array}$$



hat gleiche Normalform!
Man kann sehen: $A^1(x_0, \bullet) = (1/2, 1/2)$, $A^2(q, \bullet) = (1/2, 1/2)$ (im ersten Bsp.) ein Nash-gleichgewicht in Verhaltensstrategien, es gibt keins in reinen!

Ausblick:

1) Auch gemischte Strategien definieren eine Normalform:

   (1) Die Auszahlung ist

   $$C_M^{i\mu} = \int_{\mathfrak{S}} C_\alpha^{i\mu} \, dM(\alpha)$$

   (rein mechanisch, wie für Bimatrix-Spiele)

(2) Die Verteilung von einer gemischten Strategie induziert ist gegeben durch

$$m_\mu^M := \int_{\mathfrak{S}} m_\mu^\alpha(\bullet)\, dM(\alpha)$$

Dann wäre die Auszahlung

$$\overset{\circ}{C}{}_M^{i\mu} := \int_{\underline{\mathsf{X}}} C^i(x)\, dm_\mu^M(x) = \int_\Omega C_0^i X\, dP$$

für X nach m_μ^M verteilt.

Satz 3.4.11. *1)* $\overset{\circ}{C}{}_M^{i\mu} = C_M^{i\mu}$

2) Für jedes M existiert eine Verhaltensstrategie A mit gleicher Verteilung:

$$m_\mu^M = m_\mu^A$$

bei kontanter Infostruktur; Satz von Kuhn)

3) Also existieren Nashgleichgewichte in Verhaltensstrategie!