

Genetische Algorithmen

Michael Gerhäuser

Universität Bayreuth

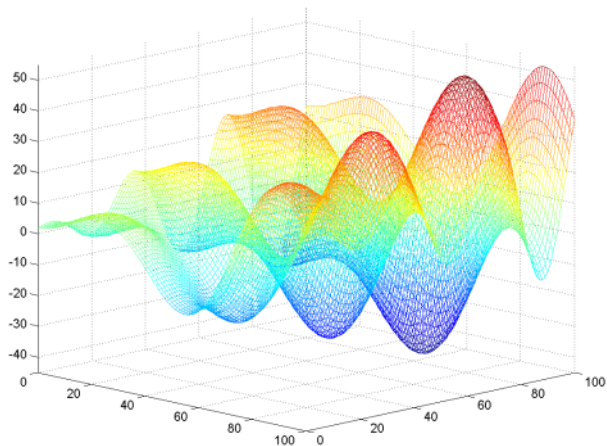
28.06.2008



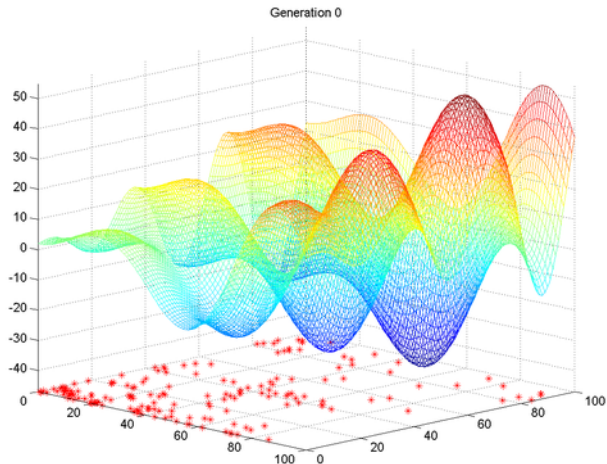
UNIVERSITÄT
BAYREUTH

- 1 Motivierende Beispiele
- 2 Grundbegriffe der biologischen Evolution
- 3 Formalisierung der Grundbegriffe
- 4 Modellierung mit Markov-Ketten
- 5 Konvergenzanalyse des Markov-Modells
- 6 No-Free-Lunch-Theorem

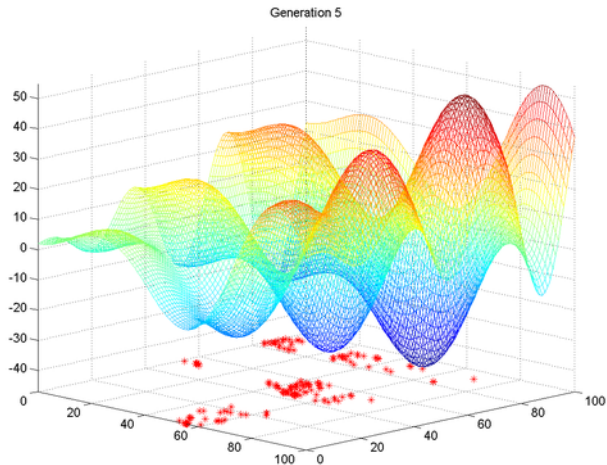
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



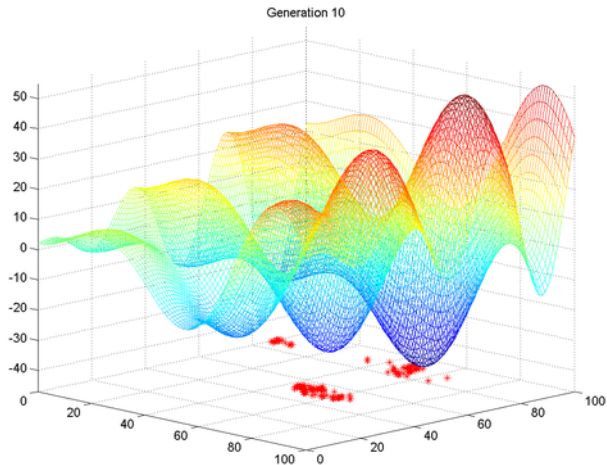
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



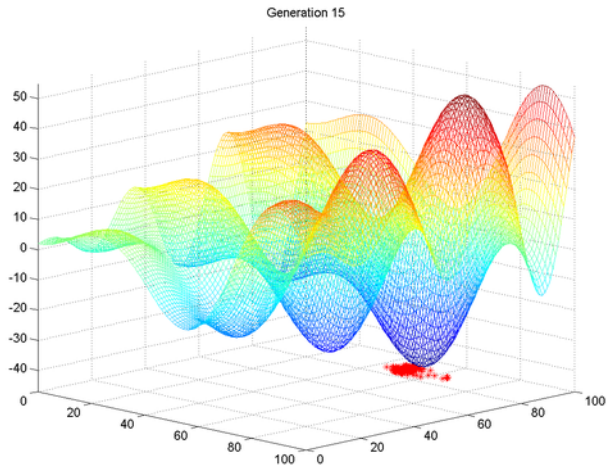
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



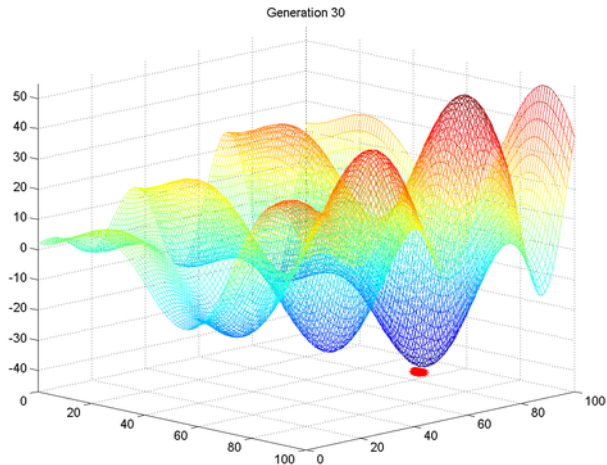
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



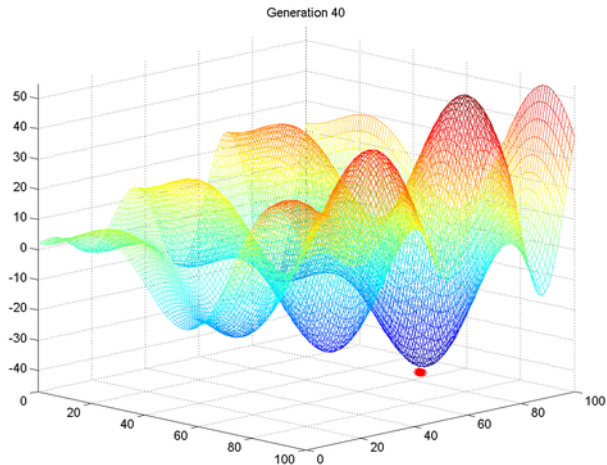
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



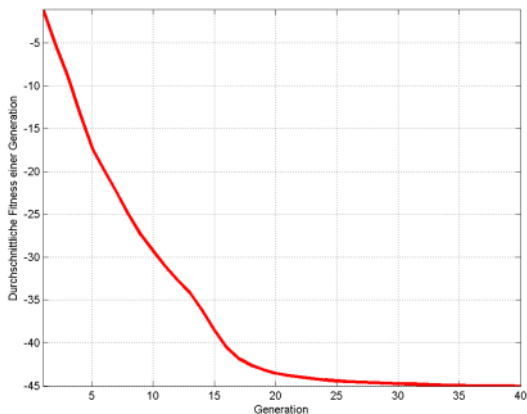
$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



$$f : D \rightarrow \mathbb{R}, (x, y) \mapsto \frac{x}{3} \cdot \sin\left(\frac{\pi}{35} \cdot x\right) + \frac{y}{3} \cdot \cos\left(\frac{2\pi}{35} \cdot y\right) + 2$$



Grundbegriffe der biologischen Evolution

- Genotyp, Phänotyp, Individuum, Population und Evolution
- Evolutionäre Operatoren

Evolution

Genotyp und Phänotyp

- Der *Phänotyp* eines Lebewesens ist sein äußeres Erscheinungsbild.
- Maßgeblich beeinflusst von seiner DNS, einem in allen Lebewesen vorkommendem Biomolekül.
- Die Grundbausteine von DNS sind *Gene*
- Das *Genom* ist die Menge aller Gene eines Lebewesens.
- Als *Genotyp* bezeichnet man den individuellen Satz von Genen eines Lebewesens.

Evolution

Individuum und Population

Die Kombination aus Genotyp und Phänotyp eines Lebewesens heißt *Individuum*. Eine Menge von Individuen nennt man *Population*.

Evolution

Evolution beschreibt die langsame, kontinuierlich fortschreitende Entwicklung und Veränderung der Lebewesen auf der Erde.

Evolutionäre Operatoren

Rekombination

Erzeugen neuer Individuen (**Kinder**) durch

- Austausch und
- Neuordnung

des genetischen Materials mehrerer bestehender Individuen (**Eltern**).

Evolutionäre Operatoren

Mutation

Veränderungen im Genom eines Individuums, z.B. aufgrund von

- Fehlern bei der Reproduktion von Genomen
- äußeren Einwirkungen wie Verstrahlung.

Mutation tritt sehr selten auf.

Evolutionäre Operatoren

Selektion

- Das Auswählen von Individuen aus einer Population zur Rekombination.
- Je höher die Fitness eines Individuum, desto wahrscheinlicher seine Auswahl.
- Durch die Selektion wird die durchschnittliche Fitness einer Population verbessert ⇒ **Optimierung**.

Wie kann man sich den Optimierungscharakter der biologischen Evolution zur Lösung von Optimierungsproblemen mit dem Computer zunutze machen?

Formalisierung der Grundbegriffe

- Allgemeines Optimierungsproblem
- Grundbegriffe der biologischen Evolution formal definiert
- Formale Definition der Evolutionsoperatoren mit Beispielen
- Evolutionärer Algorithmus

Allgemeines Optimierungsproblem

Definition (Allgemeines Optimierungsproblem)

Ein *Optimierungsproblem* $\mathfrak{P} := (\Omega, \mathcal{F}, \preceq)$ ist gegeben durch

- einen Suchraum Ω ,
- einer Zielfunktion

$$\mathcal{F} : \Omega \rightarrow \mathbb{R},$$

die jedem $x \in \Omega$ einen Güte-/Fitnesswert $\mathcal{F}(x)$ zuweist und

- einer Vergleichsoperation $\preceq \in \{\leq, \geq\}$.

Die Menge \mathcal{X} der globalen Optima ist definiert als

$$\mathcal{X} := \{x \in \Omega \mid \mathcal{F}(x) \preceq \mathcal{F}(y) \quad \forall y \in \Omega\}.$$

Individuum, Genotyp und Phänotyp

Definition (Genotyp und Phänotyp)

Analog zur biologischen Evolution setzt sich ein Individuum aus Genotyp und Phänotyp zusammen. Der Phänotyp eines Individuums ist das entsprechende x aus dem Suchraum Ω . Der Genotyp eines Individuums sei definiert über die bijektive Abbildung dec^{-1} vom Suchraum in die Menge aller möglichen Genotypen $G := G_{\Omega}$. Abhängig vom Problem kann man G relativ frei wählen.

Beispiele für Genotypen

Standard-Binärcodierung

Die intuitivste genotypische Individuumsrepräsentation. Entspricht der binären Darstellung einer Zahl im Computer. Ein Gen entspricht einem Bit im sog. **Bit-String**.

Gravierender Nachteil: Invertierung eines Bits kann zu drastischer Änderung im Phänotyp führen und umgekehrt können sich Genotypen benachbarter Phänotypen in jedem Bit unterscheiden:

3 und 4 haben die Darstellung $3 = 011$ und $4 = 100$.

Beispiele für Genotypen

Gray-Codierung

Gray-Code beseitigt das Problem nicht, reduziert es jedoch. Im Gegensatz zum Standard-Binär Code ist der Hammingabstand von $k - 1$ und k immer 1.

Vergleich Standard-Binär- und Gray-Code

Zahl k	Standard-Binär	Hammingdist.	Gray	Hammingdist.
0	000	-	000	-
1	001	1	001	1
2	010	2	011	1
3	011	1	010	1
4	100	3	110	1

Beispiele für Genotypen

Nichtbinäre Repräsentation

- Vektoren reeller Zahlen
- Permutationen

Definition (Binäre Operatoren)

Seien i und j Bit-Strings der Länge l . Dann sei

- 1 $i \oplus j$ das *logische exklusive oder* von i und j ,
- 2 $i \otimes j$ das *logische und* aus i und j ,
- 3 $|i| := \sum_{\nu=1}^l i_{\nu}$ die Anzahl der Einsen in i und
- 4 $H(i, j) := |i \oplus j|$ der *Hamming-Abstand* von i und j , also die Anzahl an unterschiedlichen Bitpositionen.

Beispiele zu Binäroperatoren

Definition (Besondere Bit-Strings)

σ_s sei der Bit-String, dessen Einträge bis zur Position s nur aus Einsen und sonst nur aus Nullen besteht und τ_s der zu σ_s komplementäre Bit-String.

Binäroperatoren

Seien $l = 5$, $i = (1, 0, 1, 0, 1)$ und $j = (1, 1, 0, 0, 0)$. Dann ist

- 1 $i \oplus j = (0, 1, 1, 0, 1)$,
- 2 $i \otimes j = (1, 0, 0, 0, 0)$,
- 3 $|i| = 3$, $|j| = 2$,
- 4 $H(i, j) = 3$,
- 5 $\sigma_2 = (1, 1, 0, 0, 0)$ und $\tau_2 = (0, 0, 1, 1, 1)$.

Definition Individuum und Population

Definition

Sei $\mathfrak{P} = (\Omega, \mathcal{F}, \preceq)$ ein Optimierungsproblem und $G := \text{dec}^{-1}(\Omega)$. Das Tupel $(x, g := \text{dec}^{-1}(x), \mathcal{F}(x)) =: I(x)$ heißt *Individuum*. $\mathcal{F}(x)$ wird *Fitness* des Individuums $I(x)$ genannt.

Weiterhin sei $\mathcal{P}(\mathbb{P})$ die Potenzmenge von

$$\mathbb{P} := \{I(x) \mid x \in \Omega\} \subset (\Omega \times \text{dec}^{-1}(\Omega) \times \mathbb{R}),$$

der Menge aller zu \mathfrak{P} gehörenden Individuen.

Ein $P \in \mathcal{P}(\mathbb{P})$ nennt man *Population*.

Definition der Evolutionsoperatoren

Definition (Mutationsoperator)

Ein Mutationsoperator ist eine Abbildung

$$M : \mathcal{P}(\mathbb{P}) \rightarrow \mathcal{P}(\mathbb{P}), P \mapsto P'$$

mit $|P| = |P'|$. Je nach Stärke der Änderungen im Phänotyp eines Individuums erfüllt der Mutationsoperator verschiedene Aufgaben:

- Große Änderungen \Rightarrow Erreichbarkeit aller Punkte im Suchraum gewährleisten, erhöht die Streuung.
- Kleine Auswirkungen \Rightarrow Mutation tastet die nähere Umgebung eines Individuums ab und übernimmt Feinabstimmung.

Beispiele für Evolutionsoperatoren

Mutationsoperator

- Invertieren eines oder mehrerer Bits.
- Uniforme Verteilung
- Verteilung mit kleinerer Wahrscheinlichkeit, je höher die Bitposition.

Definition der Evolutionsoperatoren

Definition (Rekombinationsoperator)

Ein Rekombinationsoperator mit e Eltern und k Kindern ist eine Abbildung

$$R^{e,k} : \mathcal{P}(\mathbb{P}) \rightarrow \mathcal{P}(\mathbb{P}), E \mapsto K \supset E,$$

mit $|E| = e, e + k = |K|$.

⇒ Rekombinationsoperatoren erweitern die Population. Auch hier gibt es mehrere Varianten:

- **Kombinierende Operatoren** holen Informationen nur aus den Genotypen der Eltern, ordnen vorhandenes Erbmateriale neu an, erfinden keine neuen Genbelegungen.
- **Expandierende Operatoren** nutzen zur Bestimmung einer Suchrichtung zusätzliche Informationen (Gütwerte).

Beispiele für Evolutionsoperatoren

Rekombination

1 *1-Punkt-Crossover*

Gegeben: 2 Bit-Strings a, b der Länge $l = 5, s \in \mathbb{N} : 0 \leq s \leq l$.

$$a = (0, 1, 1, 0, 1), b = (1, 1, 0, 0, 0) \Rightarrow c = (0, 1, 0, 0, 0).$$

2 *N-Punkt-Crossover*

Analog (1) an N Punkten s_1, \dots, s_N mit $s_1 < s_2 < \dots < s_N$ und $N \leq l$.

3 *Uniform Crossover*

Jedes Bit wird aus den Elternbits an gleicher Position mit uniformer Verteilung ausgewählt. Mit l, a, b aus (1) wäre $c = (1, 1, 0, 0, 1)$ ein möglicher Nachfolger.

Definition der Evolutionsoperatoren

Definition (Selektionsoperator)

Ein *Selektionsoperator* ist eine Abbildung

$$S^{e,k} : \mathcal{P}(\mathbb{P}) \rightarrow \mathcal{P}(\mathbb{P}), E \mapsto K \subset E$$

mit $|E| = e > k = |K|$.

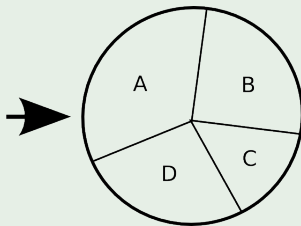
⇒ Wählt aus e Individuen k aus.

- **Probabilistische Selektion:** Auswahl eines Individuums mit einer bestimmten Wahrscheinlichkeit.
- **Deterministische Selektion:** Auswahl allein auf Basis der Fitness eines Individuums ⇒ *elitäre Selektion*.

Beispiele für Operatoren

Selektion

- Fitnessproportionale Selektion: Auswahlwahrscheinlichkeit proportional zur Fitness



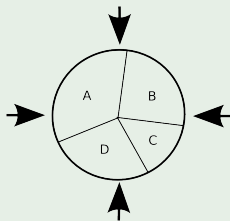
Jedes Individuum bekommt ein Kreissegment zugewiesen. Drehen der Kreisscheibe und Auswahl des Individuums auf den der Pfeil zeigt, bis benötigte Anzahl Individuen ausgewählt.

Beispiele für Operatoren

Selektion: Stochastic universal sampling

Problem: Bei Populationen mit extrem ungleichen Fitnesswerten gibt es Individuum mit sehr großem Kreissegment, so dass mit relativ großer Wahrscheinlichkeit nur dieses ausgewählt wird.

Lösung:



Beispiele für Operatoren

Wettkampfselektion

- Wähle zufällig $q \in \{2, \dots, |P|\}$ Individuen.
- Übernehme das Individuum mit dem besten Fitnesswert.
- Wiederhole, bis benötigte Anzahl Individuen ausgewählt ist.

Rangbasierte Selektion

- Deterministisch
- Es werden die besten Individuen ausgewählt.

Evolutionärer Algorithmus

Evolutionärer Algorithmus

Eingabe: Zielfunktion f , Populationsgröße μ ,

Zahl der Kinder λ

Erzeuge Population P der Größe μ

Berechne $f(i)$ für alle $i \in P$

Wiederhole

 Generiere λ Kinder durch Anwenden des

 Rekombinationsoperators auf P

 Speichere Kinder und Eltern in P'

 Wende Mutationsoperator auf P' an

 Wähle mit Selektionsoperator aus P' μ aus

 und speichere sie in P

bis Abbruchkriterien erfüllt

Ausgabe: P

Folgende Fragen sollten sich uns nun stellen:

- Konvergiert so ein Genetischer Algorithmus?
- Falls er konvergiert, wogegen konvergiert er?
- Welche Bedingungen müssen erfüllt sein, dass er konvergiert?

Welche Eigenschaften haben Genetische Algorithmen?

- Zufallsbasierter, iterativer Algorithmus
⇒ Stochastischer Prozess
- Es ist nur die aktuelle Population relevant, vorhergehende Populationen spielen keine Rolle
⇒ Markov-Kette

Modellierung mit Markov-Ketten

- Grundlagen der Stochastik
- Definition stochastischer Prozess & Markov-Kette
- Exemplarische Modellierung eines Genetischen Algorithmus als Markov-Kette

Grundlagen der Stochastik

Definition (Wahrscheinlichkeitsmaß, -raum)

- Ein Maß P auf einem messbaren Raum (Ω, \mathcal{A}) mit $P(\Omega) = 1$ heißt *Wahrscheinlichkeitsmaß*.
- Ω heißt *Zustandsraum*.
- Mengen $A \in \mathcal{A}$ heißen *Ereignisse*, $P(A)$ ist die *Wahrscheinlichkeit* des Ereignisses A .
- (Ω, \mathcal{A}, P) heißt *Wahrscheinlichkeitsraum*.

Grundlagen der Stochastik

Definition (Bedingte Wahrscheinlichkeit)

Seien $A, B \in \mathcal{A}$ Ereignisse mit $P(B) > 0$. Dann ist

$$P_B(A) := P(A | B) := \frac{P(A \cap B)}{P(B)}$$

die *bedingte Wahrscheinlichkeit* von A bezüglich B .

Grundlagen der Stochastik

Definition (Zufallsvariable)

Sei $(\Omega_1, \mathcal{A}_1, P)$ ein Wahrscheinlichkeitsraum und $(\Omega_2, \mathcal{A}_2)$ ein messbarer Raum.

- Eine $\mathcal{A}_1/\mathcal{A}_2$ -messbare Funktion $X : \Omega_1 \rightarrow \Omega_2$ heißt Zufallsvariable.
- Meist ist $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}^k, \mathbb{B}^k)$.

Beispiel

Definition (Erwartungswert)

Sei $X : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}^k, \mathbb{B}^k)$ eine Zufallsvariable. Dann heißt

$$E X := \int_{\Omega} X dP$$

der *Erwartungswert* der Zufallsvariable X .

Beispiel

Beispiel zu den stochastischen Grundbegriffen

$$\begin{aligned}\Omega &:= \{Kopf, Zahl\}, \\ \Rightarrow \mathcal{A} &= \{\emptyset, \{Kopf\}, \{Zahl\}, \{Kopf, Zahl\}\}, \\ P(A) &:= \begin{cases} 0 & \text{falls } A = \emptyset, \\ \frac{1}{2} & \text{falls } A \in \{\{Kopf\}, \{Zahl\}\}, \\ 1 & \text{falls } A = \Omega, \end{cases} \\ X(\omega) &:= \begin{cases} 0 & \text{falls } \omega = Kopf, \\ 1 & \text{falls } \omega = Zahl, \end{cases} \\ EX &= \int_{\Omega} X dP = \frac{1}{2} \sum_{i=1}^2 X(\omega_i) = \frac{1}{2}.\end{aligned}$$

Grundlagen der Stochastik

Definition

Sei X eine Zufallsvariable und $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen in (Ω, \mathcal{A}, P) . Die Folge (X_n) konvergiert gegen X

- *total* $(X_n \xrightarrow{c} X)$, wenn für jedes $\epsilon > 0$ gilt:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n P(\{\omega \in \Omega \mid |X_i(\omega) - X(\omega)| > \epsilon\}) < \infty.$$

- *fast sicher* $(X_n \xrightarrow{f.s.} X)$, wenn

$$P(\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Zusammenhänge zwischen den Konvergenzbegriffen

Definition (und Satz)

- *in Wahrscheinlichkeit* $(X_n \xrightarrow{P} X)$, wenn für jedes $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| > \epsilon\}) = 0.$$

Ist X eine Zufallsvariable und (X_n) eine Folge von Zufallsvariablen in (Ω, \mathcal{A}, P) , so gilt:

$$X_n \xrightarrow{c} X \quad \Rightarrow \quad X_n \xrightarrow{f.s.} X \quad \Rightarrow \quad X_n \xrightarrow{P} X.$$

Definition stochastischer Prozess & Markov-Kette

Definition (Stochastischer Prozess)

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und (E, \mathcal{E}) ein messbarer Raum. Eine Folge von Zufallsvariablen $X_n : (\Omega, \mathcal{A}, P) \rightarrow (E, \mathcal{E})$ heißt *stochastischer Prozess* mit *Zustandsraum* E .

Definition stochastischer Prozess & Markov-Kette

Definition (Markov-Kette)

- Ein stochastischer Prozess X_n mit Zustandsraum E heißt Markov-Kette, wenn für alle $n \geq 0$ und alle $x_0, \dots, x_{n+1} \in E$ gilt

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = \\ P(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

- Eine Markov-Kette heißt *homogen*, wenn für alle $i, j \in E$

$$P(X_{n+1} = i \mid X_n = j) =: p_{ij},$$

unabhängig von n ist.

Definition stochastischer Prozess & Markov-Kette

Definition (Übergangsmatrix)

Die durch p_{ij} definierte Matrix $\Pi \in \mathbb{R}^{|E| \times |E|}$ heißt *Übergangsmatrix* der Markov-Kette. Π ist eine stochastische Matrix, d.h.

$$p_{ij} \geq 0 \quad \forall i, j \in |E| \quad \text{und} \quad \sum_{j=1}^{|E|} p_{ij} = 1 \quad \forall i \in |E|.$$

Markov-Modell

Im folgenden Abschnitt betrachten wir einen genetischen Algorithmus zum Lösen eines Optimierungsproblems mit folgenden Eigenschaften:

- Individuen werden als Standard Bit-String der Länge l codiert.
- Die Population soll eine konstante Größe von n haben.
- Rekombination R : 1-Punkt-Crossover.

Markov-Modell

- Zu jeder Zeile von R gehört eine Population, auf die der Evolutionsoperator angewandt wird und eine Population gehört zu jeder Spalte, die nach Anwenden des entsprechenden Operators entstehen kann.
- Wir modellieren R als stochastische Matrix

Rekombination im Markov-Modell

Rekombinationsmatrix

- Ausgangspunkt ist eine Population (i_1, i_2) mit 2 Individuen.
- Durch Rekombination wird ein drittes Individuum i_3 erzeugt.
- Benötigen Wahrscheinlichkeit $P_R(\alpha, \beta, \gamma)$, mit der Individuum γ durch 1-Punkt-Crossover aus α und β entsteht.
- Hierfür zunächst eine Hilfsfunktion:

$$\mathcal{I}_{\alpha, \beta, \gamma}(s) := \begin{cases} 1, & \text{falls } \gamma = (\alpha \otimes \sigma_s) \oplus (\beta \otimes \tau_s) \\ 0, & \text{sonst} \end{cases}$$

Rekombination im Markov-Modell

Rekombinationsmatrix

Damit erhält man P_R und somit auch R einfach durch Aufsummieren über alle $l + 1$ möglichen Crossoverpunkte:

$$R_{\varphi,\psi}^{\alpha,\beta,\gamma} := \begin{cases} P_R(\alpha, \beta, \gamma) = \frac{1}{l+1} \left(\sum_{s=0}^l \mathcal{I}_{\alpha,\beta,\gamma}(s) \right) & \text{falls } \varphi = \alpha \text{ und} \\ & \psi = \beta \\ 0 & \text{sonst.} \end{cases}$$

R, M, S sind stochastisch

Theorem

Die Matrizen R, M und S sowie $\Pi := R \cdot M \cdot S$ sind stochastisch.

Beweis (nur für M und Π):

$$\sum_{(\alpha, \beta, \gamma) \in \mathbb{P}^3} R_{\lambda, \mu}^{\alpha, \beta, \gamma} = \sum_{\gamma \in \mathbb{P}} P_R(\lambda, \mu, \gamma) = 1$$

R, M, S sind stochastisch

Beweis (Fortsetzung).

Seien $A \in \mathbb{R}^{k \times l}$, $B \in \mathbb{R}^{l \times m}$ stochastische Matrizen, dann ist $C := A \cdot B$ stochastisch:

$$\begin{aligned} \sum_{j=1}^m \sum_{i=1}^l a_{ni} \cdot b_{ij} &= \sum_{i=1}^l \sum_{j=1}^m a_{ni} \cdot b_{ij} = \sum_{i=1}^l \left(a_{ni} \cdot \sum_{j=1}^m b_{ij} \right) = \\ &= \sum_{i=1}^l a_{ni} = 1. \end{aligned}$$



Konvergenzanalyse des Markov-Modells

- Hinreichende Konvergenzbedingungen
- Notwendige Konvergenzbedingungen

Hinreichende Bedingungen

Betrachte Genetische Algorithmen mit den folgenden Eigenschaften:

- Rekombination und Mutation sorgen dafür, dass von jedem Individuum aus ein optimales Individuum erreichbar ist.
- Der Selektionsoperator ist elitär.

Hinreichende Bedingungen

Nach eventueller Umsortierung der Indizes hat die zugehörige Übergangsmatrix die Gestalt

$$P = \begin{pmatrix} I_\nu & 0 \\ R & Q \end{pmatrix}.$$

- ν ist die Anzahl der optimalen Zustände und I_ν die Einheitsmatrix der Dimension ν .
- $R \in \mathbb{R}^{(N-\nu) \times \nu}$ und $Q \in \mathbb{R}^{(N-\nu) \times (N-\nu)}$ nicht negative Matrizen, derart dass für alle $i \in \{\nu + 1, \dots, N\}$ ein $t \in \mathbb{N}$ existiert mit

$$\sum_{j=1}^{\nu} p_{ij}^t > 0.$$

Hinreichende Bedingungen

Theorem (Hinreichende Konvergenzbedingung)

Ein Genetischer Algorithmus über einem endlichen Suchraum Ω , mit elitärem Selektionsoperator und dessen evolutionäre Operatoren sicherstellen, dass von jedem Individuum aus ein optimaler Zustand erreichbar ist, konvergiert unabhängig von der Startpopulation total, fast sicher und in Wahrscheinlichkeit gegen eine Optimallösung.

Hinreichende Bedingungen

Beweisidee.

Man zeigt zunächst, dass

$$\lim_{t \rightarrow \infty} \Pi^t = \lim_{t \rightarrow \infty} \begin{pmatrix} I_\nu & 0 \\ R & Q \end{pmatrix}^t = \begin{pmatrix} I_\nu & 0 \\ (I - Q)^{-1}R & 0 \end{pmatrix}.$$

Damit hat man die Konvergenz in Wahrscheinlichkeit gezeigt, für die totale Konvergenz bleibt zu zeigen:

$$\sum_{t=0}^{\infty} \sum_{i \in \{1, \dots, N\}} \sum_{j \in \{\nu+1, \dots, N\}} p_{ij}^t < \infty.$$



Notwendige Bedingungen

Hierfür werden einige Ergebnisse aus der Markov-Ketten-Theorie benötigt.

Theorem

Sei (A_i) eine Folge von Matrizen $\in \mathbb{R}^{n \times n}$ mit $\lim_{i \rightarrow \infty} A_i = A$.
Dann gilt:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A_i = A.$$

Notwendige Bedingungen

Theorem (Ergodensatz)

Ist $A \in \mathbb{R}^{n \times n}$ stochastisch, so existiert stets:

$$Q := \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i.$$

Theorem

Sei $A \in \mathbb{R}^{n \times n}$ stochastisch und irreduzibel, dann ist

$$Q = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$$

eine positive Matrix, d.h. $q_{ij} > 0 \quad \forall i, j \in \{1, \dots, n\}$.

Notwendige Bedingungen

Theorem (Notwendige Konvergenzbedingungen)

Notwendig für eine von der Startpopulation unabhängige Konvergenz eines Genetischen Algorithmus über einem endlichen Suchraum gegen ein Optimum sind

- *Rekombination und Mutation sorgen für einen Übergang von jedem Individuum zu jedem Anderen.*
- *Selektion ist elitär.*

Notwendige Bedingungen

Beweisskizze.

Wäre von einer Population i aus keine Optimale erreichbar, so ist

$$\sum_{j \in \{1, \dots, \nu\}} p_{ij}^t = 0 \quad \forall t \in \mathbb{N}.$$

Eine Konvergenz unabhängig von der Startpopulation kann also nicht gewährleistet werden.

Notwendige Bedingungen

Beweisskizze (Fortsetzung).

Wäre die Selektion nicht elitär, so besäße jede Population eine positive Überlebenswahrscheinlichkeit. Durch die Transitivität von Mutation und Rekombination ist auch die gesamte Übergangsmatrix transitiv und somit ist es möglich von jeder Population zu jeder anderen zu gelangen, insbesondere können auch optimale Populationen wieder verlassen werden. $\Pi = RMS$ ist also irreduzibel, aber natürlich auch weiterhin stochastisch.

Notwendige Bedingungen

Beweisskizze (Fortsetzung).

Falls $\lim_{t \rightarrow \infty} \Pi^t$ existiert, so ist der Grenzwert

$$\lim_{t \rightarrow \infty} \Pi^t = Q = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \Pi^i$$

eine positive Matrix. Man erhält also für jede Population i und beliebige $t \in \mathbb{N}$

$$\lim_{t \rightarrow \infty} \sum_{j \in \{1, \dots, \nu\}} p_{ij}^t < 1.$$

⇒ Der Genetische Algorithmus konvergiert nicht unbedingt gegen ein Optimum. □

Zusammenfassung: No-Free-Lunch-Theorem

No-Free-Lunch Theorem oder: Von nichts kommt nichts

Im Gegensatz zu deterministischen Algorithmen sind Genetische Algorithmen

- zufallsbasiert, jedoch
- stellen sie keine Einschränkungen an das zu lösende Problem wie Differenzierbarkeit.

Theorem (No-Free-Lunch)

Je zwei Algorithmen A_1, A_2 lösen beliebige Probleme P_1, P_2 aus der Menge aller Probleme im Durchschnitt gleich gut.

Anwendungsgebiete:

- Erstellen von Fahr-, Stunden- und Raumplänen
- Optimierung (siehe Motivation)
- Lösen NP-schwerer Probleme wie TSP, Rucksackproblem

- Grundlagen der Maßtheorie: 1
- Modellierung der Mutationsmatrix: 5
- Modellierung einer Selektionsmatrix: 8

Definition (σ -Algebra)

Ein Mengensystem \mathcal{A} mit $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt σ -Algebra, wenn

- $\Omega \in \mathcal{A}$,
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,
- $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Definition (Maßraum, Maß)

Sei \mathcal{A} eine σ -Algebra aus Teilmengen von Ω , dann heißt (Ω, \mathcal{A}) messbarer Raum oder Maßraum. Eine Abbildung

$$\mu : \mathcal{A} \rightarrow \mathbb{R}_0^+$$

heißt Maß, wenn $\mu(\emptyset) = 0$ und μ σ -additiv, d.h. sind $A_1, A_2, \dots \in \mathcal{A}$ abzählbar viele paarweise disjunkte Mengen und A ihre Vereinigungsmenge, dann gilt

$$\mu(A) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

Definition

Seien $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$ Maßräume. Eine Abbildung

$$X : \Omega_1 \rightarrow \Omega_2$$

heißt $\mathcal{A}_1/\mathcal{A}_2$ -messbar, wenn für alle $A_2 \in \mathcal{A}_2$ gilt: $X^{-1}(A_2) \in \mathcal{A}_1$.

Zurück zum Vortrag: 1

Mutation

Mutationsmatrix

- Ausgangspunkt ist eine Population (i_1, i_2, i_3) mit 3 Individuen.
- Mutation wirkt nur auf i_3 .
- Benötigen Wahrscheinlichkeit $P_M(\alpha, \beta)$, mit der aus Individuum α durch Mutation β wird.
- Ist p_M die Wahrscheinlichkeit, dass ein Bit mutiert und $H(\alpha, \beta) = l - k$, so gilt:

$$P_M(\alpha, \beta) = p_M^{l-k} \cdot (1 - p_M)^k.$$

Mutation

Mutationsmatrix

Damit lässt sich die Mutationsmatrix M schreiben als:

$$M_{\alpha,\beta,\gamma}^{\lambda,\mu,\nu} := \begin{cases} P_M(\gamma, \nu) & \text{falls } \lambda = \alpha \text{ und } \mu = \beta \\ 0 & \text{sonst.} \end{cases}$$

Beweis, dass M stochastisch.

Zu jedem Individuum α gibt es $\binom{l}{j}$ Individuen β mit $H(\alpha, \beta) = j$.
Damit gilt mit Summation über alle j :

$$\sum_{\beta \in \mathbb{P}} P_M(\alpha, \beta) = \sum_{j=0}^l \binom{l}{j} p_M^j \cdot (1 - p_M)^{l-j} = 1.$$

$$\Rightarrow \sum_{(\alpha, \beta, \gamma) \in \mathbb{P}^3} M_{\lambda, \mu, \nu}^{\alpha, \beta, \gamma} = \sum_{\gamma \in \mathbb{P}} P_M(\nu, \gamma) = 1.$$



Zurück zum Vortrag: 6

Selektion

Selektionsmatrix

- Auch hier gehen wir von einer Population (i_1, i_2, i_3) mit 3 Individuen aus.
- Selektion wählt 2 der Individuen “mit zurücklegen” über fitnessproportionale Selektion aus.
- Benötigen Wahrscheinlichkeit $P_S(\alpha, \beta, \gamma, \lambda)$, mit der Individuum λ aus (α, β, γ) ausgewählt wird:

$$P_S(\alpha, \beta, \gamma, \lambda) = \begin{cases} \frac{f(\lambda)}{f(\alpha)+f(\beta)+f(\gamma)}, & \text{falls } \lambda \in \{\alpha, \beta, \gamma\} \\ 0, & \text{sonst.} \end{cases}$$

Selektion

Selektionsmatrix

Da die beiden Individuen unabhängig voneinander ausgewählt werden, ist die Matrix S definiert durch

$$S_{\alpha, \beta, \gamma}^{\lambda, \mu} := P_S(\alpha, \beta, \gamma, \lambda) \cdot P_S(\alpha, \beta, \gamma, \mu).$$

Beweis, dass S stochastisch.

$$\begin{aligned}\sum_{\alpha \in \mathbb{P}} \sum_{\beta \in \mathbb{P}} S_{\lambda, \mu, \nu}^{\alpha, \beta} &= \sum_{\alpha \in \mathbb{P}} \sum_{\beta \in \mathbb{P}} P_S(\lambda, \mu, \nu, \alpha) \cdot P_S(\lambda, \mu, \nu, \beta) = \\ &= \sum_{\alpha \in \mathbb{P}} \left(P_S(\lambda, \mu, \nu, \alpha) \cdot \sum_{\beta \in \mathbb{P}} P_S(\lambda, \mu, \nu, \beta) \right) = \\ &= \sum_{\alpha \in \mathbb{P}} P_S(\lambda, \mu, \nu, \alpha) = 1.\end{aligned}$$



Zurück zum Vortrag: 6