Wintersemester 2005/06

Fundamente der Computational Intelligence

(Vorlesung)

Prof. Dr. Günter Rudolph

Fachbereich Informatik

Lehrstuhl für Algorithm Engineering

Inhalt

- Grundlagen Optimierung
- Nachbarschaftssuche
- (1+1)-EA
    - Erwartete Laufzeit
        - Counting Ones
        - Leading Ones
    - Konvergenz
- Populationsbasierte EA
    - EA im $\mathbb{B}^n$
    - EA im $\mathbb{R}^n$

Nachtrag: EP-Selektion (Beispiel für Verlust des Besten)

$\mu = \lambda = 2$ bei Minimierung, $q = 2$

Index	1	2	3	4
f(x)	12.3	1.9	4.2	8.1
Rang	4	1	2	3
Siege	0	2	2	2
Sortierrang	4	3	2	1
Auswahl	-	-	+	+

Annahme: Sortierung nach Anzahl der Siege mit „selection sort“

$\Rightarrow$ abhängig von Implementierung!

$\Rightarrow$ Lösung: Bestem **vorher** q+1 Siege zuweisen!

Schematischer Ablauf eines Evolutionären Algorithmus

	GA	ES	EP
`Initialisierung der Population`	x	x	x
`repeat`			
`  Reproduktionsselektion`	x	-	-
`  Rekombination`	x	x	-
`  Mutation`	x	x	x
`  Überlebensselektion`	-	x	x
`until Abbruch`			

Heute: Keine Unterscheidung zwischen GA/ES/EP, da Übergänge fließend

EA im $\mathbb{R}^n$

Selektion: unabhängig von Representation

Rekombination: k-Punkt-Crossover, Uniform Crossover
funktioniert auf allen Produkträumen!

zusätzlich:
Intermediäre Rekombination $z_i = \alpha_i x_i + (1-\alpha_i) y_i$, $\alpha_i \in (0,1)$

$\alpha_i \equiv \alpha$: auf Verbindungslinie zwischen x und y
sonst innerhalb des durch x und y definierten Hyperrechtecks

auch: Varianten mit > 2 Eltern

Mutation: additiv via Normalverteilung N(0, C), C Kovarianzmatrix
häufig: $C = \sigma^2 \cdot I_n$ oder individuelle σ_i

auch: Cauchyverteilung („Levy flights“) oder Mixturen

notwendig: Anpassung von C bzw. σ_i während Suche!

Optimierung ohne „Anpassen“ (reine Zufallssuche)

$f(x) = \| x \|^2 = x'x \to \min!$ wobei $x \in S_n(r) = \{ x \in \mathbb{R}^n : \| x \| \leq r \}$

Z_k ist gleichverteilt in $S_n(r)$

$X_{k+1} = Z_k$ falls $f(Z_k) < f(X_k)$, sonst $X_{k+1} = X_k$

$\Rightarrow V_k = \min \{ f(Z_1), f(Z_2), \ldots, f(Z_k) \}$ bester ZF-Wert bis Iteration k

$P\{ \| Z \| \leq x \} = P\{ Z \in S_n(x) \} = \mathrm{Vol}(S_n(x)) / \mathrm{Vol}(S_n(r)) = (x / r)^n \ , \ 0 \leq x \leq r$

$P\{ \| Z \|^2 \leq x \} = P\{ \| Z \| \leq x^{1/2} \} = x^{n/2} / r^n \ , \ 0 \leq x \leq r^2$

$P\{ V_k \leq x \} = 1 - (1 - P\{ \| Z \|^2 \leq x \})^k = 1 - (1 - x^{n/2} / r^n)^k$

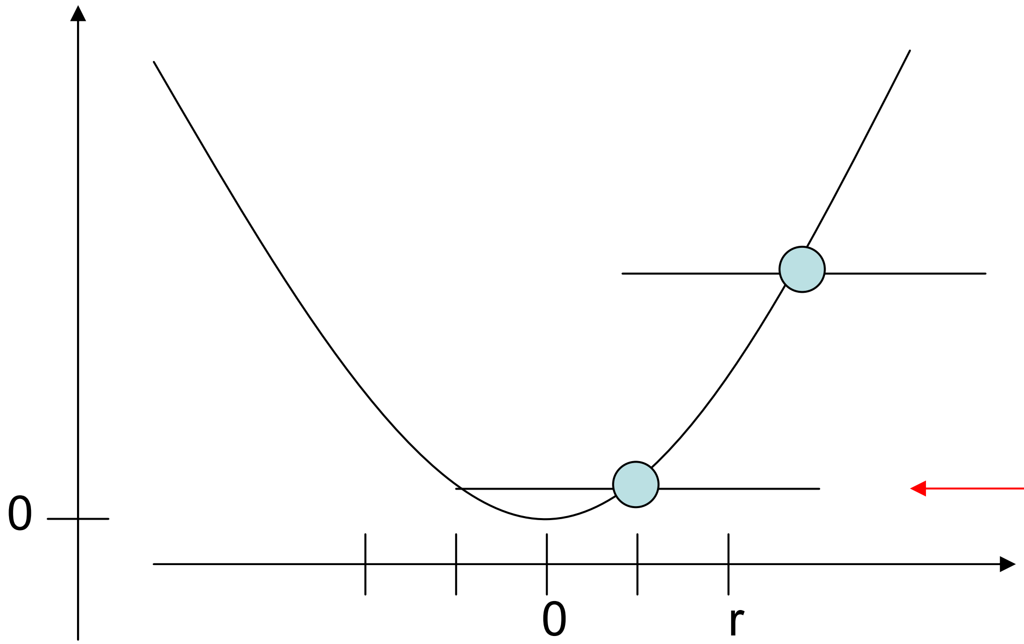
$E[V_k] \to r^2 \, \Gamma(1 + 2/n) \, k^{-2/n}$ für große k

Optimierung ohne „Anpassen“ (lokal gleichverteilt)

$f(x) = \| x \|^2 = x'x \rightarrow \min!$ wobei $x \in S_n(r) = \{ x \in \mathbb{R}^n : \| x \| \leq r \}$

Z_k ist gleichverteilt in $[-r, r]$, $n = 1$

$X_{k+1} = X_k + Z_k$ falls $f(X_k + Z_k) < f(X_k)$, sonst $X_{k+1} = X_k$



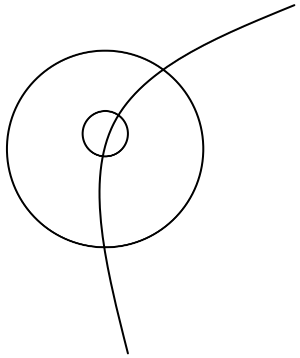
ab hier:
wie reine
Zufallssuche!

(1+1)-EA mit Schrittweitenanpassung (1/5-Erfolgsregel, Rechenberg 1973)

Idee:

- Wenn viele erfolgreiche Mutationen, dann Schrittweite zu klein.
- Wenn wenige erfolgreiche Mutationen, dann Schrittweite zu groß.

bei infinitesimal kleinem Radius ist Erfolgsrate = 1/2



Ansatz:

- Protokolliere erfolgreiche Mutationen in gewissem Zeitraum
- Wenn Anteil größer als gewisse Schranke (z. B. 1/5), dann Schrittweite erhöhen, sonst Schrittweite verringern

Schrittweite?

Normalverteilter Vektor $Z = N(0, \sigma^2 I_n)$ hat zufällige Länge $|| Z ||$

Man kann zeigen: $E[\, || Z || \,] \approx \sigma \, (n - ½)^{½}$ mit Varianz von etwa $\sigma^2 / 2$

Definition:

Zufallsvektor $X = (x_1, \ldots, x_n)$ sphärisch verteilt
wenn X für jede orthonormale Matrix T
die gleiche Verteilung wie TX besitzt (also rotationsinvariant ist). ■

Satz:

$X = (x_1, \ldots, x_n)$ sphärisch verteilt gdw. $X = r \cdot U$, wobei

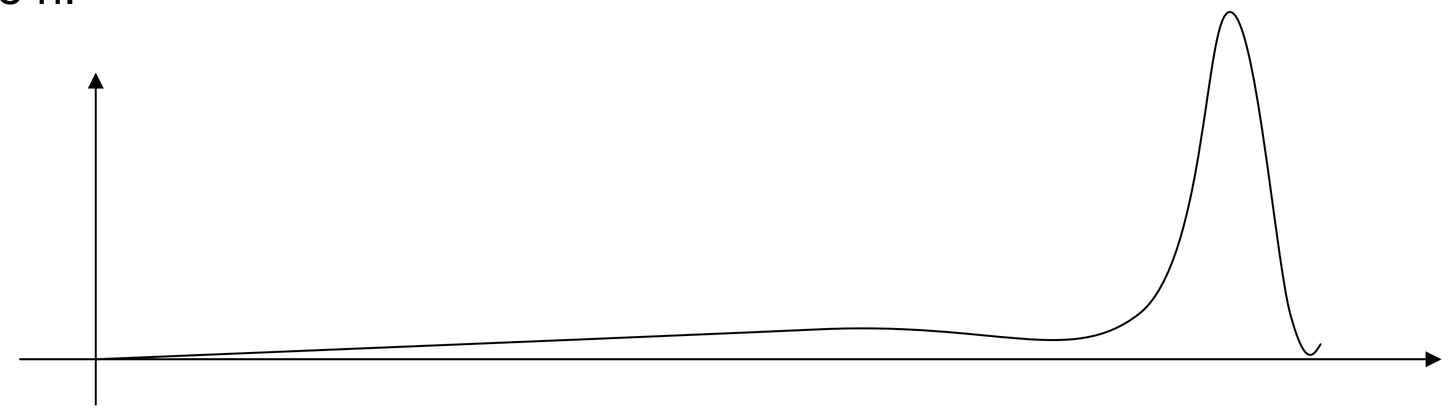
r nicht negative Zufallsvariable,

U n-dim. Zufallsvektor gleichverteilt auf Kugelrand und

= im Sinne von „gleich in Verteilung“ ■

Wenn Zufallsschrittweite r wie $\chi_n(\sigma)$-verteilt, dann X normalverteilt $N(0, \sigma^2 I_n)$

Für große n:



$\Rightarrow$ Normalverteilung verhält sich ähnlich wie Kugelrandverteilung!

$\Rightarrow$ fast alle sphärischen Verteilungen verhalten sich für große n ähnlich!

fast alle? alle mit finiter Varianz!

Analyse

(1, λ)-EA mit $f(x) = \| x \|^2$

$$\| Y_k \|^2 = \| X_k + r_k U_k \|^2 = (X_k + r_k U_k)' (X_k + r_k U_k)$$

$$= X'_k X_k + 2 r_k X'_k U_k + r_k^2 U'_k U_k$$

$$= \|X_k\|^2 + 2 r_k X'_k U_k + r_k^2 \underbrace{\|U_k\|^2}_{=1} = \|X_k\|^2 + 2X'_k U_k + r_k^2$$

da das zufällige Skalarprodukt x‘U die gleiche Verteilung hat wie ||x|| B,

wobei Zufallsvariable B betaverteilt mit Parametern (n-1)/2 auf [-1, 1] ist, folgt

$$\| Y_k \|2 = \| X_k \|^2 + 2 r_k \| X_k \| B + r_k^2$$

da der (1,λ) den besten Wert aus λ Versuchen selektiert, folgt

$$\| X_{k+1} \|^2 = \| X_k \|^2 + 2 r_k \| X_k \| B_{1:\lambda} + r_k^2$$

$$\| X_{k+1} \|^2 = \| X_k \|^2 + 2r_k \| X_k \| B_{1:\lambda} + r_k^2$$

$\downarrow$ bedingte Erwartungswerte auf beiden Seiten

$$E\| X_{k+1} \|^2 = \| X_k \|^2 + 2r_k \| X_k \| E[B_{1:\lambda}] + r_k^2$$

$\downarrow$ Ansatz: $r_k = \gamma \| X_k \|$

$$E\| X_{k+1} \|^2 = \| X_k \|^2 + 2 \gamma \| X_k \|^2 E[B_{1:\lambda}] + \gamma^2 \| X_k \|^2$$

$\downarrow$ wg. Symmetrie von B folgt $E[B_{1:\lambda}] = - E[B_{\lambda:\lambda}] < 0$

$$E\| X_{k+1} \|^2 = \| X_k \|^2 - 2 \gamma \| X_k \|^2 E[B_{\lambda:\lambda}] + \gamma^2 \| X_k \|^2$$
$$= \| X_k \|^2 (1 – 2\gamma E[B_{\lambda:\lambda}] + \gamma^2)$$

Parabelscheitel bei $\gamma^* = E[B_{\lambda:\lambda}]$, also $E\| X_{k+1} \|^2 = \| X_k \|^2 (1 – E[B_{\lambda:\lambda}]^2)$