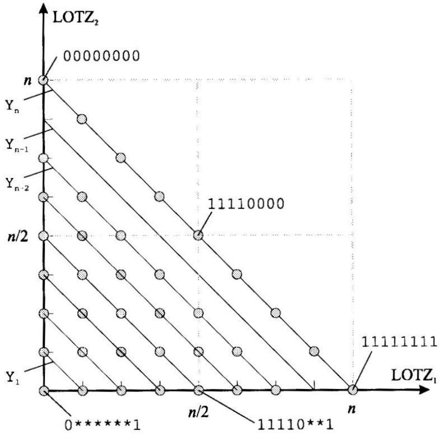
Praktische Optimierung

Günter Rudolph
Fakultät für Informatik, Lehrstuhl XI
Fachgebiet *Computational Intelligence*

21-JAN-2008

Laufzeitanalyse: LOTZ $\to$ max!

$$f(x) = \left(\sum_{i=1}^{n} \prod_{j=1}^{i} x_j, \sum_{i=1}^{n} \prod_{j=i}^{n} (1 - x_j) \right) , \quad x \in \{0,1\}^n$$



Entnommen aus Laumanns (2003), S. 60.

Laufzeitanalyse: Algorithmus SEMO

wähle $x \in \mathcal{X}$ zufällig
$P_0 = \{x\}$, setze $k = 0$
`repeat`
  wähle $x \in P_k$ zufällig
  erzeuge Nachkommen x' durch 1-Bit-Mutation aus x
  falls $\nexists y \in P_k : y \prec x' \vee f(y) = f(x')$
    $P_{k+1} = (P_k \setminus \{z \in P_k : x' \prec z\}) \cup \{x'\}$
  $k{+}{+}$
`until` Stoppkriterium erfüllt

Satz (Laumanns 2003)
Erwartete Laufzeit von SEMO auf LOTZ ist $\Theta(n^3)$.

Laufzeitanalyse: Algorithmus FEMO

wähle $x \in \mathcal{X}$ zufällig, setze Zähler $c(x) = 0$

$P_0 = \{x\}$, setze $k = 0$

```
repeat
```

wähle $x \in P_k$ mit kleinstem c-Wert; $c(x) + +$

erzeuge Nachkommen x' durch 1-Bit-Mutation aus x

falls $\nexists y \in P_k : y \prec x' \lor f(y) = f(x')$

$P_{k+1} = (P_k \setminus \{z \in P_k : x' \prec z\}) \cup \{x'\}$

c(x')=0

$k + +$

`until` Stoppkriterium erfüllt

Satz (Laumanns 2003)

Erwartete Laufzeit von FEMO auf LOTZ ist $\Theta(n^2 \log n)$.

Laufzeitanalyse: Algorithmus XYZ

wähle $x \in \mathcal{X}$ zufällig, setze $y = z = x$, $t_y = t_z = \emptyset$

`repeat`

erzeuge x', y', z' durch 1-Bit-Mutation aus x, y, z

falls $x \preceq x'$ setze $x = y = z = x'$, lösche Tabulisten

sonst akzeptiere y', z' falls $y' \parallel y$ bzw. $z' \parallel z$

und nicht Vorgänger war (Tabuliste)

`until` Stoppkriterium erfüllt

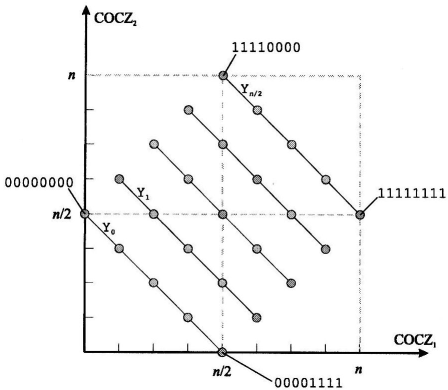
Tabulisten t_y, t_z enthalten Paretomenge

Satz (Rudolph, jetzt)

Erwartete Laufzeit von XYZ auf LOTZ ist $\Theta(n^2)$.

Laufzeitanalyse: COCZ $\rightarrow$ max!

$$f(x) = \left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n/2} x_i + \sum_{j=n/2+1}^{n} (1 - x_i) \right) \ , \quad x \in \{0,1\}^n$$



Entnommen aus Laumanns (2003), S. 62.

Laufzeitanalysen für COCZ

Satz (Laumanns 2003)
Erwartete Laufzeit von SEMO und FEMO auf COCZ ist $O(n^2 \log n)$. $\square$

Satz (Rudolph, jetzt)
Erwartete Laufzeit von XYZ auf COCZ ist $O(n^2)$. $\square$

XYZ funktioniert nur gut wenn Paretomenge ein 1-Bit-Pfad!

Bi-kriterielles Kugelmodel

$m \geq 2$ Zielfunktionen:

$$f(x) = (f_1(x), \ldots, f_m(x))' \longrightarrow \min!$$

$\longrightarrow$ Pareto-optimale Lösungen:

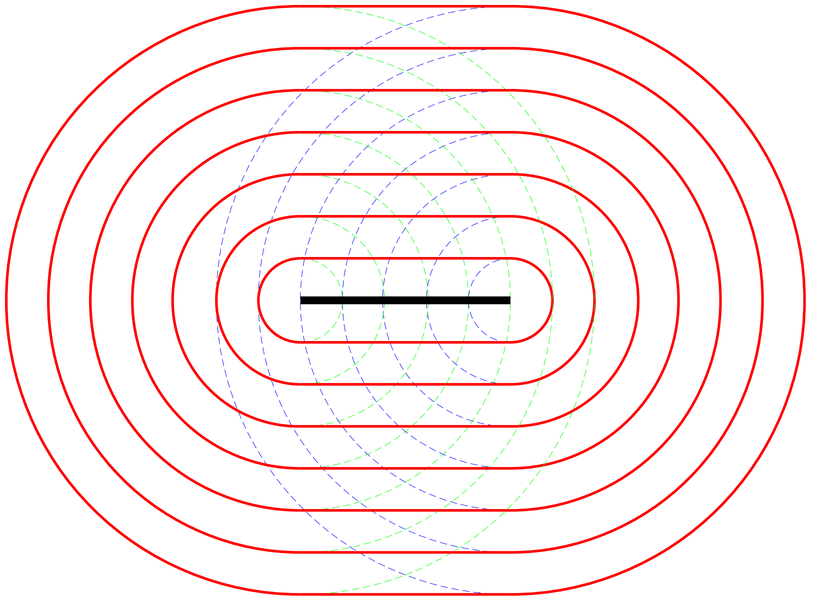
$$\nexists f(x) \in \mathbb{R}^m : (\forall i : f_i(x) \leq f_i(x^*)) \wedge f(x) \neq f(x^*)$$

Kriterium: Konvergenz des EA zur Paretomenge

hier: $f(x) = (\|x\|^2, \|x - z\|^2)'$

$$x, z \in \mathbb{R}^2; z \neq 0 \in \mathbb{R}^2$$

Höhenlinien



Algorithmus

(1) Erzeuge $X_0 \in \mathbb{R}^2$ zufällig; setze $k = 0$
(2) $Y_k = X_k + s(J, X_k) \cdot Z$
(3) if $f_J(Y_k) < f_J(X_k)$
    then $X_{k+1} = Y_k$ else $X_{k+1} = X_k$
(4) $k{+}{+}$; goto (2) solange Stoppkriterium nicht erfüllt

Z: Zufallsvektor mit E$[\,Z\,] = 0$
J: Zufallsvariable mit P$\{\,J{=}1\,\}{=}$P$\{\,J{=}2\,\}{=}\frac{1}{2}$
$s(\cdot) > 0$: Schrittweitensteuerung

Ziel: $d(X_k, \mathcal{X}^*) = \min\{\|X_k - x\| : x \in \mathcal{X}^*\} \to 0$

Numerische Vorstudien

(A) feste Schrittweiten

$$Z \sim N(0, I) \;\longrightarrow\; s(J, X_k) = \sigma > 0$$
$$Z \sim U(\partial B_\ell) \;\longrightarrow\; s(J, X_k) = r > 0$$

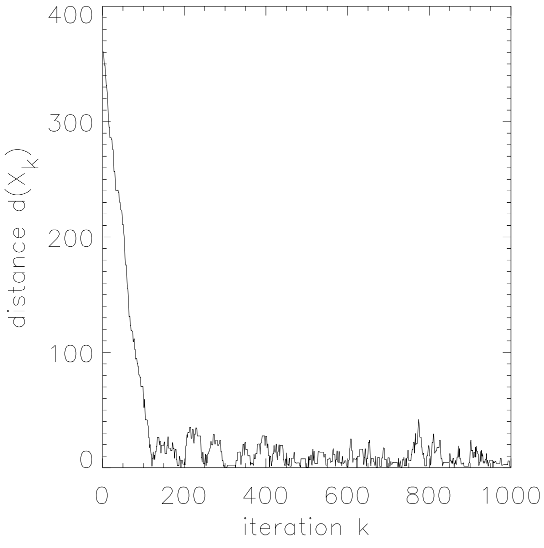
(B) opt. Schrittweite für 1 Ziel (J bekannt)

$$s(J, X_k) \;=\; 0.269\, \|\nabla f_J(X_k)\|$$
$$s(J, X_k) \;=\; 0.394\, \|\nabla f_J(X_k)\|$$

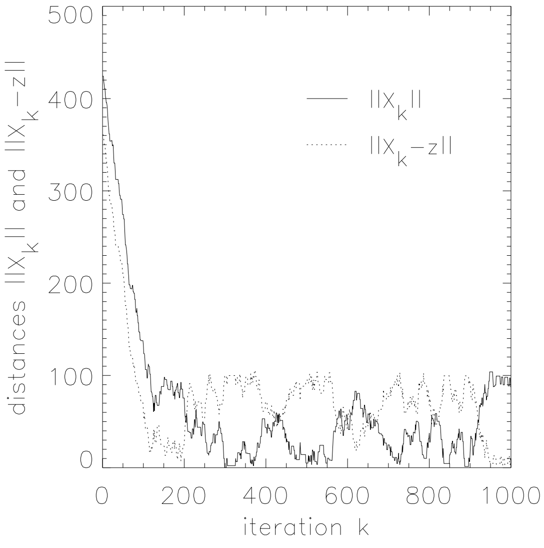
(C) proportional zum Abstand zur Paretomenge

$$s(J, X_k) = c \cdot d(X_k, \mathcal{X}^*) \qquad (c = 1)$$

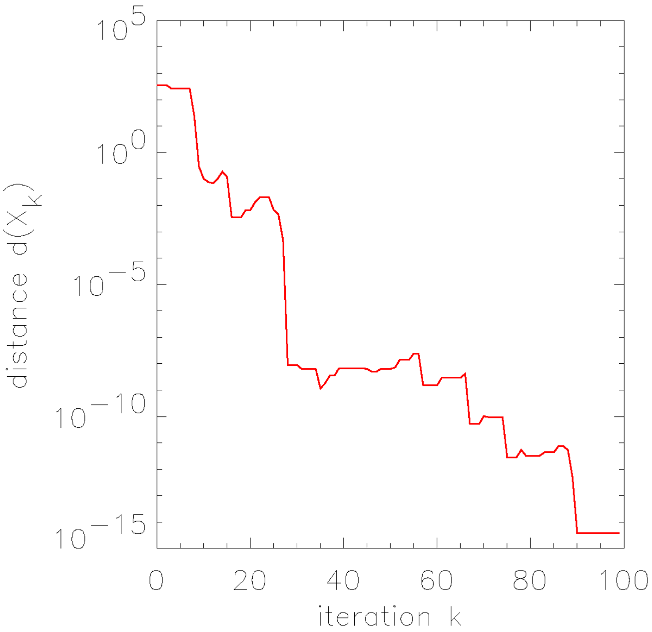
Numerische Vorstudie (A)



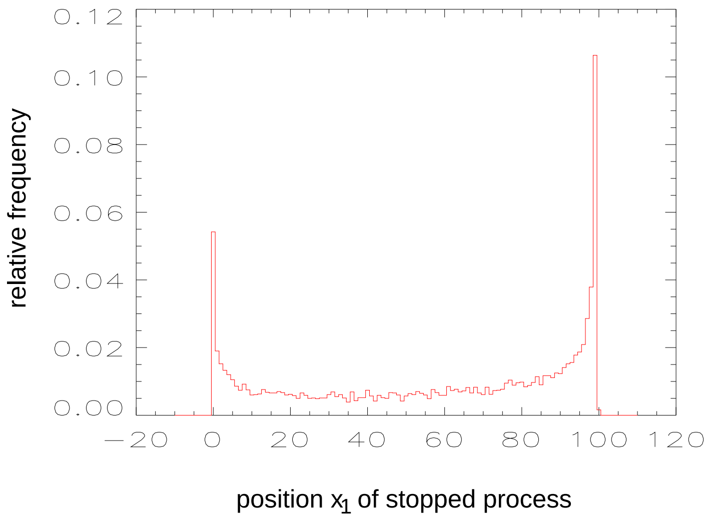
Numerische Vorstudie (A)



Numerische Vorstudie (C)



Numerische Vorstudie (C)



Mathematisches Hilfsmittel

Theorem: (Bucy & Joseph 1968)

(1) $\forall k \geq 0 : D_k \geq 0$

(2) $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ stetig, $\gamma(x) = 0 \Leftrightarrow x = 0$

(3) $\forall k \geq 0 :$ E$[\, D_k \,] < \infty$

Falls E$[\, D_{k+1} \,\|\, \mathcal{A}_k \,] \leq D_k - \gamma(D_k)$

dann $D_k \to 0$ mit W'keit 1 für $k \to \infty$ $\square$

Hier: $D_k = d(X_k, \mathcal{X}^*)$

Analyse

Finde Funktion $\gamma(\cdot)$, so dass für alle $\mathbf{x} \in \mathbb{R}^2$

$$\mathsf{E}[\, d(\mathbf{X}_{k+1}) \parallel \mathbf{X}_k = \mathbf{x}\,] \leq d(\mathbf{x}) - \gamma(d(\mathbf{x}))$$

Verteilung von $d(\mathbf{X}_{k+1})$ abh. von $\mathbf{x}_k$ sowie Zva. J and ω.
$\longrightarrow$ Stoch. Effekt von J eliminierbar durch Konditionierung

$$\begin{aligned}\mathsf{E}[\, D_{k+1} \parallel \mathbf{X}_k = \mathbf{x}\,] \quad &= \\ &\mathsf{E}[\, D_{k+1} \parallel \mathbf{X}_k = \mathbf{x}, J = 0\,] \times \mathsf{P}\{\, J = 0\,\} \quad + \\ &\mathsf{E}[\, D_{k+1} \parallel \mathbf{X}_k = \mathbf{x}, J = 1\,] \times \mathsf{P}\{\, J = 1\,\}\,.\end{aligned} \tag{1}$$

Analyse

Weitere Vereinfachung durch Ausnutzung der Symmetrien:
Da für alle $\mathbf{x} \in \mathbb{R}^2$ und $z_1 > 0$

$$\mathsf{E}[\, D_{t+1} \,\|\, \mathbf{X}_t = (x_1, x_2), J = 1\,] = \mathsf{E}[\, D_{t+1} \,\|\, \mathbf{X}_t = (z_1 - x_1, x_2), J = 0\,] \tag{2}$$

Analyse beschränkbar auf Selektion bzgl. Ziel f_0.

Hinreichend: Betrachte Fall $x_2 \geq 0$ da für alle $\mathbf{x} \in \mathbb{R}^2$ gilt

$$\mathsf{E}[\, D_{t+1} \,\|\, \mathbf{X}_t = (x_1, x_2)\,] = \mathsf{E}[\, D_{t+1} \,\|\, \mathbf{X}_t = (x_1, -x_2)\,]\,.$$

Also bestimme nur:

$$\mathsf{E}[\, D_{k+1} \,\|\, \mathbf{X}_k = \mathbf{x}, J = 0\,] = \mathsf{E}_0[\, D_{k+1}\,]$$

für alle $\mathbf{x}$ mit $x_1 \in \mathbb{R}$ und $x_2 > 0$.

Analyse

Berechne $q(\mathbf{x}) = \dfrac{\mathsf{E}_0[\,d(\mathbf{x} + s\,\mathbf{U})\,]}{d(\mathbf{x})} = \int \cdots$

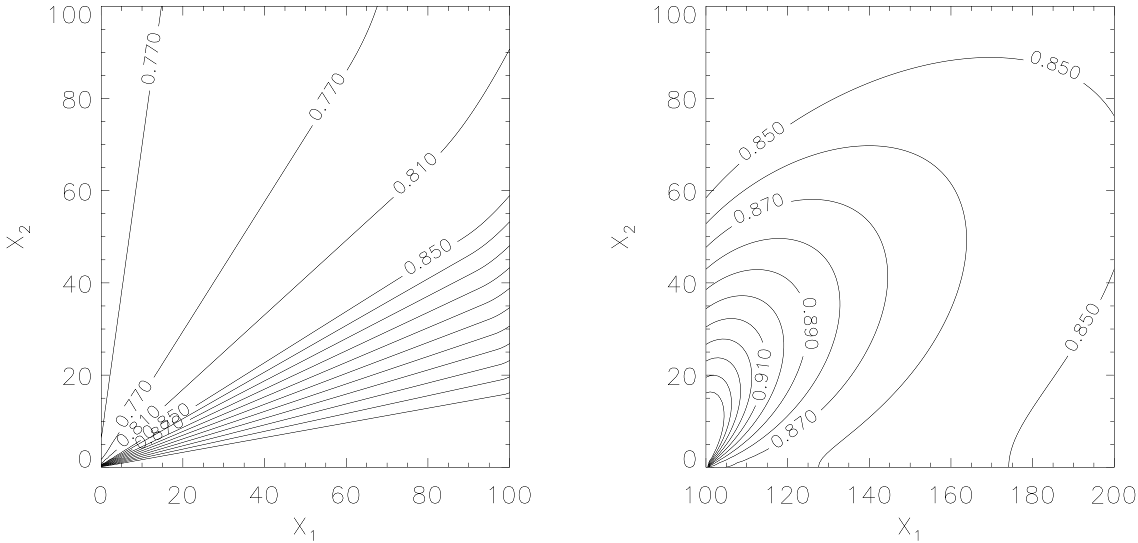
Bis jetzt: Modellierung

Ab hier: Integralrechnung (vgl. Rudolph 1998)

Resultat: Entweder $q(x) \in (0,1)$ oder $\gamma(d(x)) \to 0$

$\Rightarrow$ Konvergenz zur Paretomenge!

Konvergenz: q(x)



Konvergenzgeschwindigkeit

Fixiere $\varepsilon > 0$:

Dann $\exists c_\varepsilon \in (0,1)$ mit E$[\, D_{k+1} \,\|\, D_k \,] \leq c_\varepsilon D_k$ für alle $k > 0$ solange $D_k \geq \varepsilon$.

Problem: Abstand zur Paretomenge?

$$d(\mathbf{x}) = \frac{1}{2} \min_{0 \leq \chi \leq 1} \{ \, \| \, (1-\chi) \, \nabla f_0(\mathbf{x}) + \chi \, \nabla f_1(\mathbf{x}) \, \| \, \} \, .$$