Auf dem Weg zu einer konstruktiven Theorie von Wünschbarkeitsindizes

Dipl.-Stat. Detlef Steuer
Universität Dortmund
LS Computergestützte Statistik
steuer@gigamain.statistik.uni-dortmund.de

29. Januar 1998

- Problemstellung
- Ansätze und Anforderungen
- Wünschbarkeitsansatz
- Beispiele für Wünschbarkeiten mit einigen Eigenschaften
- Offene Fragen

Problemstellung und Ziel

Gegeben ein Objekt (Produkt) P, beschrieben durch die Eigenschaften $Y_i(P), i = 1, \ldots, Z$.
Kurz: $P = (Y_1, \ldots, Y_Z) \in I\!R^Z$.

Eigenschaften Y_i nicht direkt einstellbar!

Stattdessen $Y_i = f_i(X_1, \ldots, X_F, \epsilon_i), \quad E(\epsilon_i) = 0$.

Ziel: Gegeben mehrere Produkte P_k, Erstellen einer Rangordnung $P_{(k)}$
Implizit: Rangordnung der Einstellungen $X_k = (X_1, \ldots, X_F)_k$!

Über Modellbildung überabzählbar viele Kandidaten!

1-dim: Kein Problem Ordnungsrelation "$>$" (gelesen "besser").

mehr als 1-dim: Es existiert keine vollständige Anordnung des $I\!R^n$ für $n > 1$!

Teilausweg: Dominanz

Def: Gegeben zwei Produkte p_a und p_b, beschrieben durch dieselben Eigenschaften $Y_i(P), i = 1, \ldots, Z$. Man sag p_a dominiert p_b ($p_a > p_b$), wenn gilt: $\forall \quad i \in \{1, \ldots, Z\}$: $y_i(p_a) > y_i(p_b)$.

Nur im Falle der Dominanz ist eine objektive (problemunspezifische!) Entscheidung möglich!

Problem: Dominanz wird mit zunehmender Dimension immer seltener!

Gesucht ist dimensionsreduzierendes Funktional $C : I\!R^Z \to I\!R$ welches dann in $I\!R$ die Anordnung der Objekte ermöglicht.

Dies ist das Ausgangsproblem der multikriteriellen Optimierung (MCO, MCDM, MCDA).

Weitere Hoffnung: Nach der Dimensionsreduktion auf der Y-Seite wird evtl. Dimensonsreduktion auf der X-Seite möglich, da nur noch eine Zielgröße betrachtet wird.

Ansätze und Anforderungen

Bisherige Ansätze in Operations Research: Heuristiken

Keine Untersuchungen der statistischen Eigenschaften

- Pareto Optimalität (Problem mit hohen Dimensionen)
- Grafische Verfahren (überlagerte Plots) (Problem mit hohen Dimensionen)
- Prometee/Elektre (Problem mit hohen Dimensionen)
- Wünschbarkeitsfunktionen und -indizes
- (Monetäre) Verlustfunktionen (Probleme in der Formulierung)

Problemunspezifisch bei Dimension > 1 maximal erreichbar: Pareto-Optimalität.

Def: Gegeben eine Menge M von Produkten. Ein Produkt $p \in M$ heißt pareto-optimal in M, wenn es kein Produkt $q \in M$ gibt, das p dominiert.

Bsp: Bei Testeinkäufen erhält man für einen gewissen Geldbetrag verschiedene Mengen Äpfel und Birnen. Welches Geschäft ist pareto-optimal?

	Äpfel	Birnen
Geschäft 1	1500 g	2.25 kg
Geschäft 2	1400 g	2.35 kg
Geschäft 3	1750 g	2.4 kg

- Bei zunehmender Dimensionszahl wird Pareto-Optimalität immer seltener!
- Pareto-Optimalität nicht eindeutig!

Erste Anforderungen:

1. Das gesuchte Funktional C sollte ein Produkt aus der Menge der pareto-optimalen Produkte als Optimum bestimmen!

2. Das gesuchte Funktional C muß problemspezifisch formulierbar sein!

3. Das gesuchte Funktional C sollte größtmögliche Eindeutigkeit herbeiführen.

Wünschbarkeitsansatz

Idee 1: Übersetzen der Y_i in einheitenlose Masszahlen (Wünschbarkeiten), dann Kombination der individuellen Masszahlen zu einer Produktgütemasszahl (Wünschbarkeitsindex)

Idee 2: Güte W des Produkts P eine Funktion C der Eigenschaften $Y_i(P), i = 1, \ldots, Z$. $W(P) = C(Y_1, \ldots, Y_Z)$.

Def: Wünschbarkeitsfunktion (WF)
w : Wertebereich Y $\rightarrow [0,1], Y \hookrightarrow w(Y)$.

Wünschbarkeitsfunktionen können/sollen von problemspezifischen Parametern abhängen:
LSL_i (untere Spezifikationsgrenze),
USL_i (obere Spezifikationsgrenze),
T_i (Sollwert),
β_{li} (Gewicht für Abweichungen nach unten),
β_{ui} (Gewicht für Abweichungen nach oben)

Weitere Anforderung bei einem Sollwertproblem: monoton steigend auf $(-\infty, T_i)$, monoton fallend auf (T_i, ∞). (Garantiert Pareto-Optimalität!)

Bis hierher: Skalierung/Vergleichbarkeit Etwa: 1 Kilo Äpfel und 2500 Gramm Birnen entsprechen 0.2 $(w_1(y_1))$ bzw. 0.5 $(w_2(y_2))$

Noch keine Dimensionsreduktion!

Def: Ein Wünschbarkeitsindex (WI) W ist eine Abbildung von $[0,1]^Z \to [0,1]$.

W kombiniert w_i zu **einer** Zahl:

$$W(P) := W(w_i(Y_i)_{i=1,\ldots,Z})$$

$$= W(w_i(f_i(X_1, X_2, \ldots, X_F, \epsilon_i))_{i=1,\ldots,Z})$$

Als Funktionale W sind vor allem gebräuchlich:

- Produkt (geometrisches Mittel)
- max min $w_i(Y_i)$

Drastische Dimensionsreduktion!

Die Realisierungen der ZV P werden 1-dim bewertet und sind deshalb anordbar.

Viele Freiheiten:

- Welche w_i? Und gegeben eine Funktionenklasse für die w_i, wie werden die problemspezifischen Parameter gewählt? Wie kann ein Anwender die Parameter interpretieren?

- Welche f_i? In unserer Sprache: Welche Modelle werden für die Y_i unterstellt? Fehlerverteilung?

- Welche C? Sind Verteilungsaussagen möglich? Ist C analytisch zu behandeln?

Beispiele für Wünschbarkeitsfunktionen

Erfinder: Harrington 1965

Für Sollwertproblem Funktionen von Typ

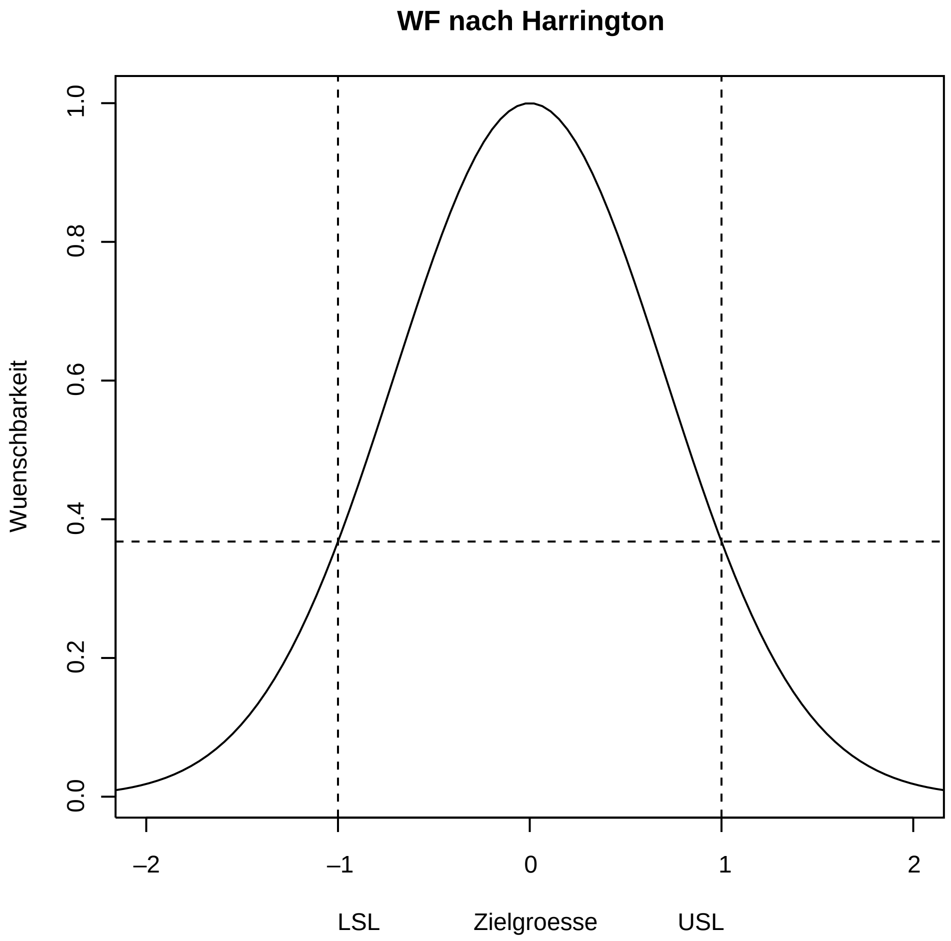
$$w_i(Y_i) = \; e^{-|T_i - Y_i|^n}.$$

Spezifikationsgrenzen haben nicht Wünschbarkeit 0, sondern Parameter n soll so gewählt werden, so daß $w(USL) = w(LSL) = 1/e$.

$$W_H(P) := \sqrt[Z]{\prod_1^Z w_i(Y_i)}$$

$$= \sqrt[Z]{\prod_1^Z w_i(f_i(X_1, X_2, \ldots, X_F, \epsilon_i))}$$

Für die Optimierungsanwendung spielt die Z-te Wurzel keine Rolle. Sie dient lediglich als Interpretationshilfe.



Nicht sehr flexibel, z.Bsp. keine nichtsymmetrischen Spezifikationen. Andere Funktionen für Maximierungsprobleme.
Differenzierbar!
Von der Statistikseite aus angehbar!

Verteilungsaussagen für WI von Harrington-Typ

ϵ_i seien i.i.d. $\sim N(0, \sigma)$

Dann $Y_i \sim N(f_i(X_1, \ldots, X_F), \sigma)$

Definiere $\mu_i := f_i(X_1, \ldots, X_F, 0)$.

Wünschbarkeitsfunktionen von Harrington:

$$w_i(Y_i) = e^{-(Y_i - T_i)^2} \text{und}$$

$$W(Y_1, Y_2, \ldots, Y_Z) = \prod_{i=1}^{Z} w_i(Y_i)$$

Gilt nun $\mu_i = T_i$, dann $(Y_i - T_i)^2 \sim \sigma\chi_1^2$.

Entsprechend

$\log W(Y_1, Y_2, \ldots, Y_Z) =$
$-\sum_1^Z (Y_i - T_i)^2 \sim -\sigma\chi_Z^2$.

Der negative Logarithmus des WI ist $\sigma\chi_Z^2$ verteilt.

Entsprechend für μ_i belibig:

Wenn $Y \sim N(\mu, \sigma^2)$
dann $(Y - T_Y)^2 \sim \sigma\chi^2_{1,(\mu-T_Y)^2/\sigma^2}$

Damit $-\log WI \sim \sigma\chi^2_{Z,\sum(\mu_i-T_{Y_i})^2/\sigma^2}$.

Einschränkung: Annahme der gleichen Varianz. Übergang zu standardisierten Größen mit Varianz $\sigma^2 = 1$.

Dann:
$-\log WI \sim \chi^2_{Z,\sum(\mu_i-T_{Y_i})^2}$ mit $\mu_i := f_i(X_1, \ldots, X_F)$

Der WI als Funktion der $X_1, \ldots, X_F$ ist *für jede Einstellung* $\mathbf{X} = (X_1, \ldots, x_F)$ nichtzentral χ^2 verteilt.

Die Varianz von -log WI ist also im Optimum minimal.

Die Minimierung des Nichtzentralitätsparameters ist für den Fall eines Linearen Modells äquivalent zu KQ-Lösung.

Im Fall linearer Modelle für die Y_i ergibt die Summation der Modelle evtl. weitere Dimensionsreduktion, diesmal auf der X-Seite!

$$Y_i = \alpha \cdot X_k \ldots, Y_j = -\alpha \cdot X_k + \ldots \Rightarrow W = 0 * X_k + ...$$

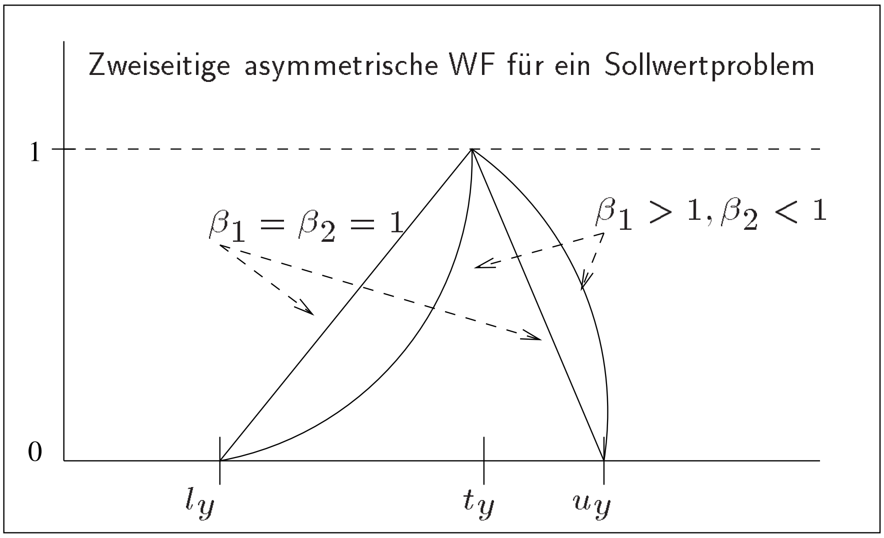
Das Auftreten der Summen hängt natürlich an der Eigenschaft der e-Funktion Produkte in Summen zu überführen. Also kann man bei einem geometrischen Mittel nur dann auf so einfache Ausdrücke kommen, wenn man WFen von Harrington-Typ kombiniert.

Zusammenhang zur Theorie der stabilen Verteilungen: Multiplikationsstabile Verteilungen sind im wesentlichen die summationsstabilen Verteilungen als Argument der e-Funktion.

Probleme: Standardisieren ist nicht immer gewünscht. Dies sind die einfachsten möglichen Annahmen. Es ist nicht zu sehen, wie man noch exakte Verteilungen für etwas kompliziertere Fälle (andere WFen) herleiten kann. (Außer Simulationen.)
Ergebnis beantwortet die falsche Frage: eigentlich geht es darum die optimale Einstellung **X** zu prognostizieren.

WFen von Derringer und Suich 1980

$$w_i(Y_i) = \begin{cases} 0, & \text{für} \quad Y_i < l_i \\ (\frac{Y_i - l_i}{t_i - l_i})^{\beta_{il}}, & \text{für} \quad l_i \leq Y_i \leq t_i \\ (\frac{u_i - Y_i}{u_i - t_i})^{\beta_{iu}}, & \text{für} \quad t_i < Y_i \leq u_i \\ 0, & \text{für} \quad u_i < Y_i \end{cases}$$



Der WI wird ebenfalls als Geometrisches Mittel gebildet.
Die β_i erlauben die Modellierung verschiedener Bedeutungen der Abweichungen von einem Sollwert für verschiedene Zielgrößen.

Vorteile: Sehr flexible Funktionenklasse
leicht auf mehrere Sprungstellen erweiterbar
Hohe Anschaulichkeit der Parameter (Scheinanschaulichkeit?)

Nachteil: Fallunterscheidungen!
– Analytische Handhabung wird kompliziert
– Abschnittsweise Definition des WI
– Sehr komplexe Modelle für den WI:

Seien $Y_i = \alpha_{0i} + \sum_{k=1}^{F} \alpha_{ki} X_k$ und $Y_j = \alpha_{0j} + \sum_{k=1}^{F} \alpha_{kj} X_k$

Dann: $Y_i \cdot Y_j = (\alpha 0 i + \sum_{k=1}^{F} \alpha_{ki} X_k) \cdot (\alpha_{0j} + \sum_{k=1}^{F} \alpha_{kj} X_k)$

Alle Wechselwirkungen treten auf!
Dieses ist der einfachste Fall $\beta_l = \beta_u = 1$ und nur für jeweils einen Abschnitt des WI!

Allg: Grad des Modells für W (in einem Abschnitt) = Summe der Grade der Individuellen w_i.

Dimensionsaufblähung!

Eindeutigkeit des Maximums noch unklar!

Fragen und Probleme

Gibt es eine Klasse von Funktionen, insbesondere geschlossen angebbar, die die Unsymmetrie der Derringer/Suich Funktionen nachbildet, differnzierbar ist und noch Verteilungsaussagen ermöglicht?

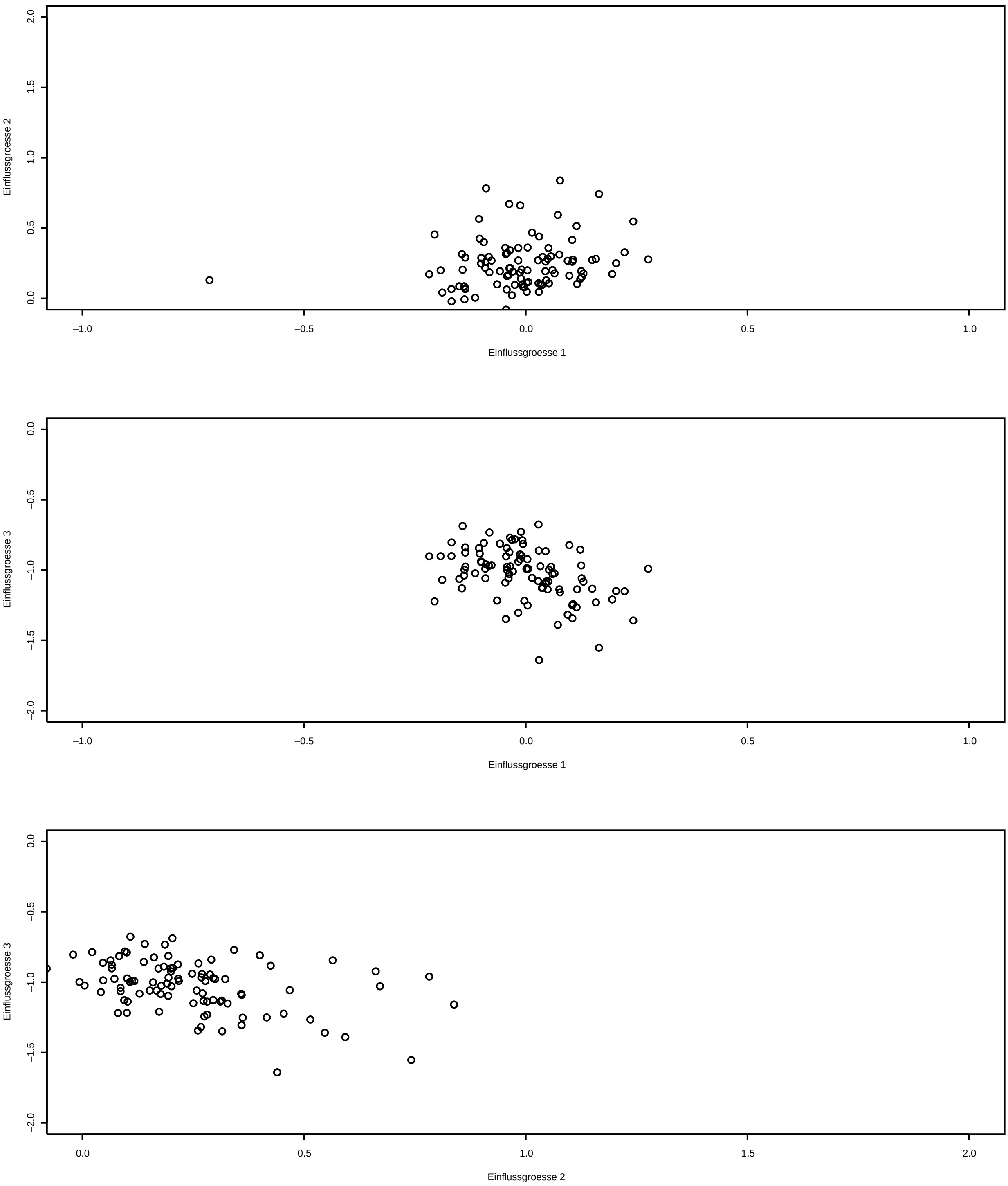
Kann man den Schritt der Modellbildung für die einzelnen Y_i überspringen und direkt den WI modellieren? Was für Modelle?
Problem bei Derringer/Suich: In welchen Abschnitt befinde ich mich?
Probleme mit den Daten: An vielen Designpunkten WI($\mathbf{X}$) $= 0$.

Wenn dies möglich ist, kann man auf Versuchspläne für WIs hoffen?

Wie sieht es mit der Varianz der Prognose der Optimalen Einstellung der X_i aus!

Wie mißt man die Güte einer solchen Prognose?

Prognostizierte Optima



Idealtool: Sebstdiagnostizierender WI Entdecken unangemessener Wfen für z.Bsp. Zielgrößen mit großer Varianz

Verteilungsaussagen über die Lage des Optimums scheinen schwer, da immer ein Optimierungsschritt involviert ist.

Auswirkung von Ausreißern auf das prognostizierte Optimum?

Fazit

Viele interessante Fragen!

Aber die komplizierte Struktur der Wünschbarkeitsindizes und des Optimierungsproblems lassen statistische Aussagen schwer möglich erscheinen.

Vielleicht ist der WI nur eine weitere Heuristik, die der statistischen Analyse nur schwer zugänglich ist?