

Optimierung

Zielsetzungen:

- Systematische Sichtweise
- Verschiedene Strategien
- Werkzeuge, aber keine Rezepte

Analyse

Im ersten Schritt der Analyse eines Problems müssen möglichst alle Inputs und Outputs gefunden werden



z.B. HPLC:

Inputs: Säulenmaterial, pH, Elutionsmittel, Gradienten, Temperatur

Outputs: Retentionszeit, Form und Fläche der Signale

z.B. Optimierung einer Reaktion

Input: Konzentrationen, Temperatur, Druck, Katalysator, ...

Output: Ausbeute

z.B. Ernährung und Gesundheit

Faktoren und irrelevante Inputs

Faktoren sind Inputs, die den Output beeinflussen

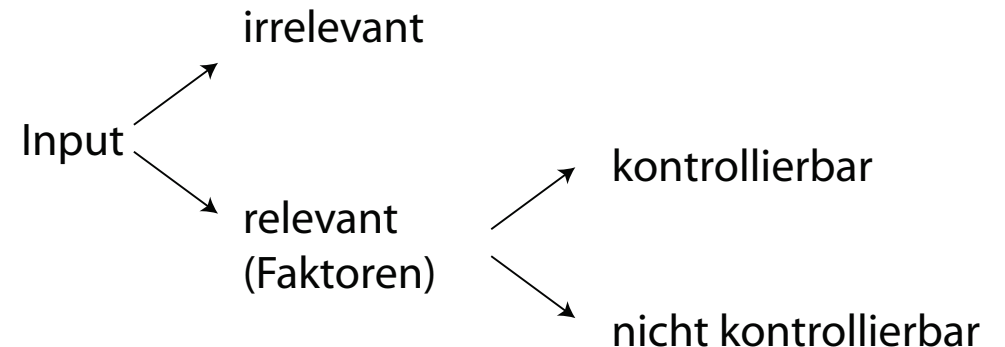


Auffinden relevanter Faktoren

Verhindern unkontrollierter Einwirkungen, z.B. Korrelation der Zeit mit Konzentrationen, Temperatur, Sonneneinstrahlung ...

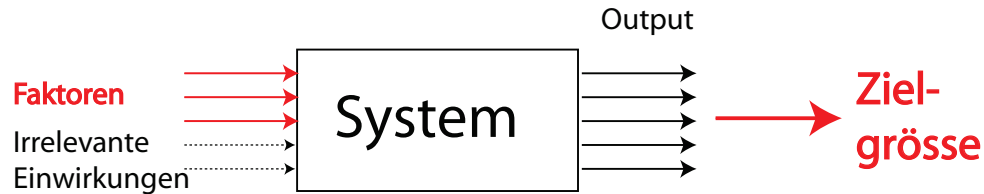
Mögliche Massnahme: Randomisierung in Zeit und/oder Raum

Kontrollierbare und nicht kontrollierbare Faktoren



Definition einer Zielgrösse

Eine Voraussetzung der Optimierung ist die Definition einer einzigen Zielgrösse



z. B. chromatographische Trennung:
Output: Chromatogramm (Retentionszeiten, Signalformen)
Zielgrösse: Eine "hinreichend gute Trennung", d.h. die Auflösung der beiden am wenigsten gut aufgelösten Signale

Vorbereitung der Optimierung

VOR der Optimierung müssen:

- eine Zielgrösse definiert werden
- der Bereich der Variablen definiert werden
- die Anzahl der Versuche festgelegt werden
- die Strategie festgelegt werden

Optimierungsmethoden

- Modellierung der Antwortfläche:
Anpassung eines linearen Modells
- Direkte Methoden:
Simplex
Variation eines Faktors pro Schritt
- Erste Ableitungen:
Box-Wilson-Methode
- Stochastische Optimierungsmethoden:
Simulated Annealing
Genetische Algorithmen

Wahl der Optimierungsmethode

Je nach Problem und Vorinformation kann die eine oder andere Methode vorteilhaft sein.

Suche des globalen Optimums:

Keine Methode ist wesentlich effizienter als das systematische Absuchen des ganzen Raumes.

Keine Methode garantiert daher das Auffinden des globalen Optimums

Stochastische Methoden sind vorteilhaft, um das globale Optimum zu suchen.

Die Modellierung der Antwortfläche ist vorteilhaft für die Feinoptimierung.

Modellierung der Antwortfläche

1. Anpassung des allgemeinen linearen Modells zweiten Grades:

z.B. für zwei unabhängige Variablen:

$$y_i = b_0 + b_1x_{1i} + b_2x_{2i} + b_{11}x_{1i}^2 + b_{22}x_{2i}^2 + b_{12}x_{1i}x_{2i} + e_i$$

2. Erste Ableitungen null setzen:

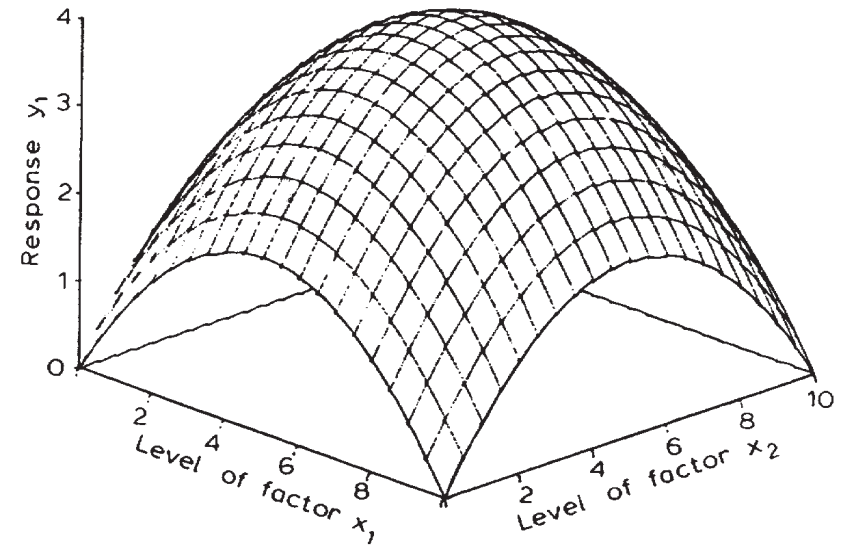
$$\text{z. B. } 0 = dy/dx_1 = b_1 + 2b_{11}x_1 + b_{12}x_2$$

$$0 = dy/dx_2 = b_2 + 2b_{22}x_2 + b_{12}x_1$$

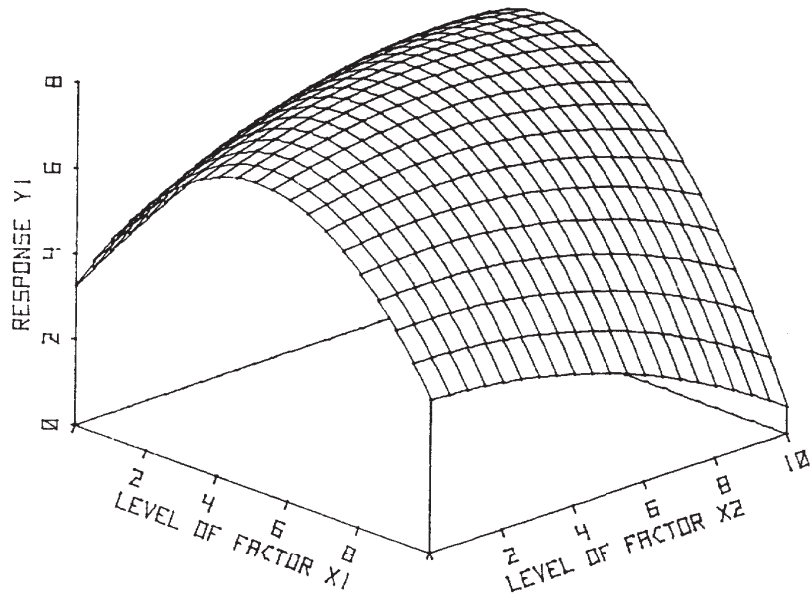
und für x_1 und x_2 lösen.

3. Sicherstellen, dass alle zweiten Ableitungen das gleiche Vorzeichen haben, positiv für ein Minimum.

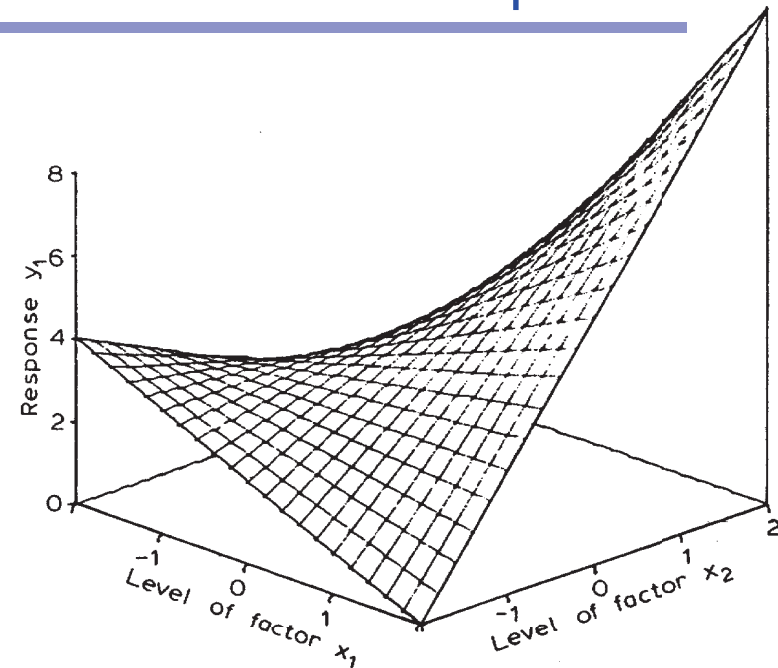
Modell ohne Wechselwirkungsterm



Modell mit Wechselwirkungsterm



Hyperfläche mit Sattelpunkt



Anzahl Parameter

Faktoren, n	Parameter, $(n+1)(n+2)/2$
2	6*
3	10
4	15
5	21

*z. B. $b_0, b_1, b_{11}, b_2, b_{22}, b_{12}$

Anzahl notwendiger Versuche

Die Anzahl anzupassender Parameter entspricht der minimalen Anzahl der notwendigen Versuche

Die Kombination der unabhängigen Variablen bei einem Optimierungsschritt muss vernünftig gewählt werden.

Zusätzliche Versuche sind nötig, wenn die Messfehler abgeschätzt und der "lack of fit" geprüft werden sollen.

--> Versuchsplanung

Anzahl notwendiger Versuche

Bei
n Messungen
f Faktorkombinationen (Kombinationen der unabhängigen Variablen)
p Modellparametern

resultieren

n-f Freiheitsgrade für die Schätzung des Messfehlers

f-p Freiheitsgrade für die Prüfung der Modellqualität

Versuchsplanung

Für die Festlegung der Faktorkombinationen eignen sich besonders:

Faktorieller Plan

Zentraler Plan

Zentral zusammengesetzter Plan

"D-optimaler" Plan

Faktorieller Plan

Anzahl Stufen (verschiedene Werte der Faktoren): k

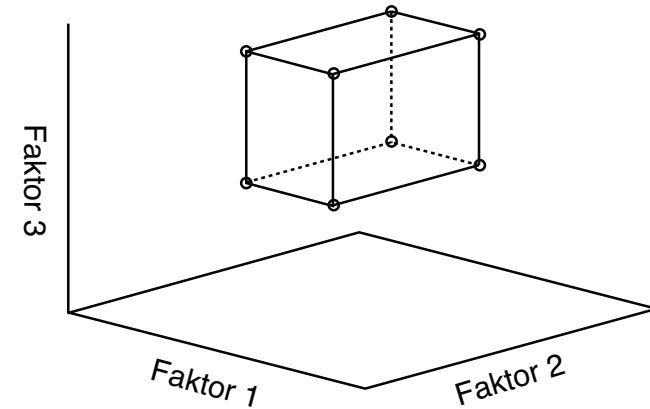
Anzahl Faktoren: n

Anzahl Versuche: k^n (wird als k^n -Plan bezeichnet)

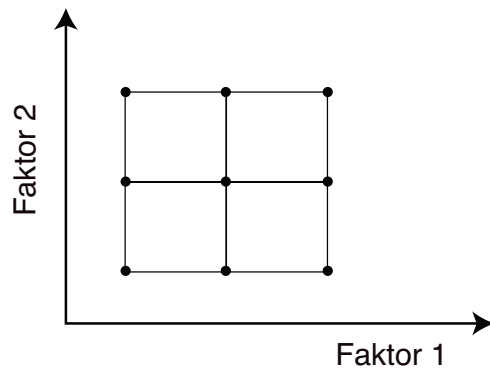
Faktoren n	Stufen k	Versuche k^n
2	2	4
2	3	9
3	2	8
3	3	27
4	2	16
4	3	81

Für ein quadratisches Modell müssen mindestens 3 Stufen vorliegen.

2^3 -Faktorieller Plan

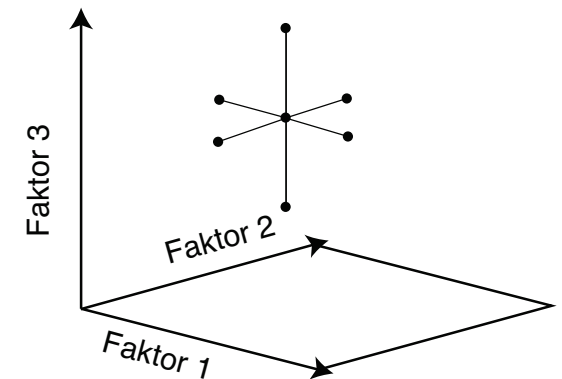
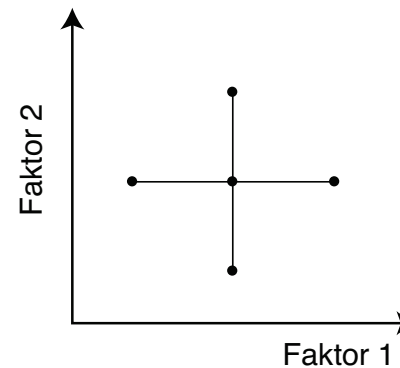


3^2 -Faktorieller Plan



Zentraler Plan

Geeignet für die Anpassung von Modellen zweiten Grades ohne Wechselwirkungsterm



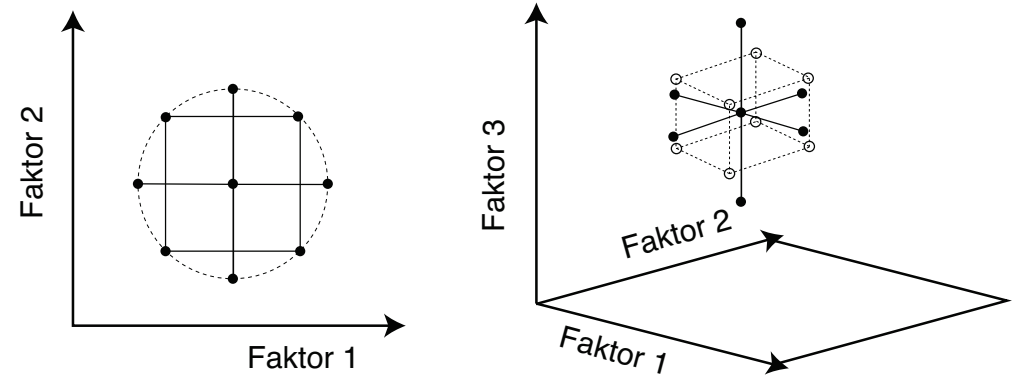
Zentraler Plan: Anzahl Versuche

Modell zweiten Grades ohne Wechselwirkungsterme

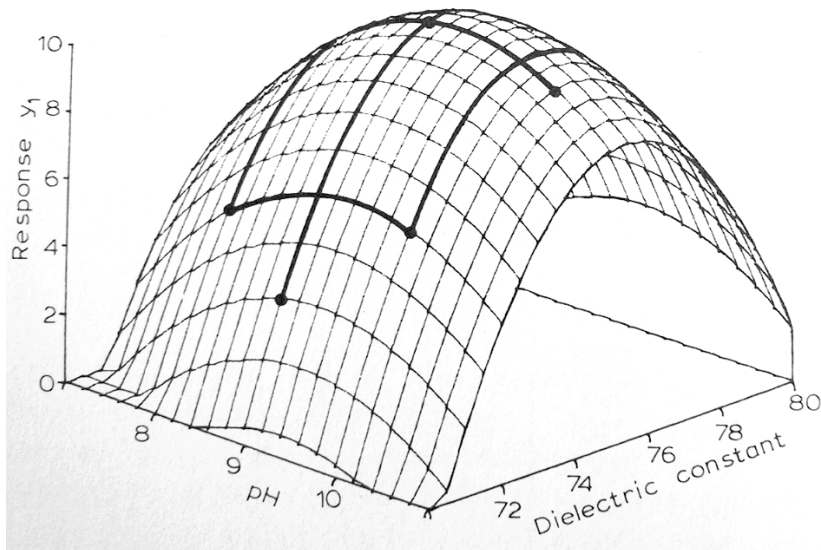
Faktoren n	Versuche $2n + 1$
1	3
2	5
3	7
4	9

Zentral zusammengesetzter Plan

Kombination eines zentralen Plans mit einem 2^n -faktoriellen Plan für die Anpassung von Modellen zweiten Grades mit Wechselwirkungstermen



Zentral zusammengesetzter Plan



Zentral zusammengesetzter Plan

Modell zweiten Grades mit Wechselwirkungstermen

Faktoren n	Versuche $2^n + 2n + 1$	Parameter $(n+1)(n+2)/2$
2	9	6
3	15	10
4	25	15

D-Optimaler Plan

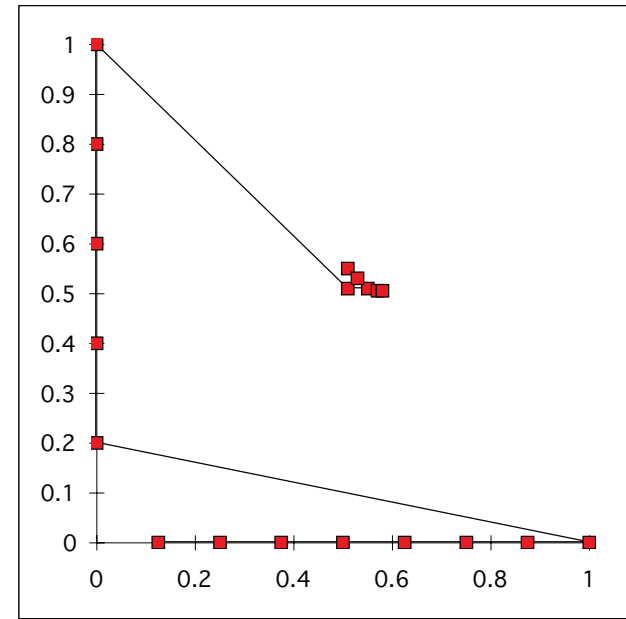
Ein Versuchsplan mit einer gegebenen Anzahl Experimente ist D-optimal, wenn die Determinante der Varianz-Kovarianzmatrix der Parameter ($\text{Det}(\mathbf{P}^T\mathbf{P})^{-1}$) minimal ist.

Dies ist gleichbedeutend mit dem Kriterium, dass die Determinante der "Informationsmatrix" ($\mathbf{P}^T\mathbf{P}$) maximal ist, da:

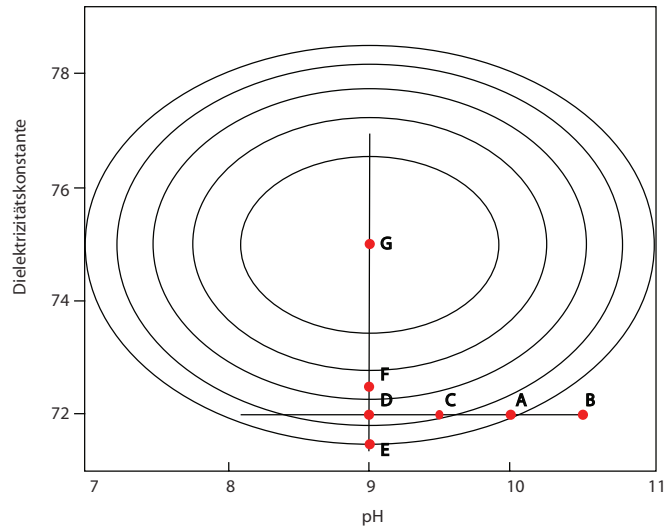
$$\text{Det}(\mathbf{P}^T\mathbf{P}) = 1/\text{Det}(\mathbf{P}^T\mathbf{P})^{-1}$$

Die oben diskutierten Pläne sind meist D-optimal. Der hier gezeigte Zusammenhang erlaubt es aber, Pläne mit beliebiger Anzahl von Versuchen optimal zu gestalten.

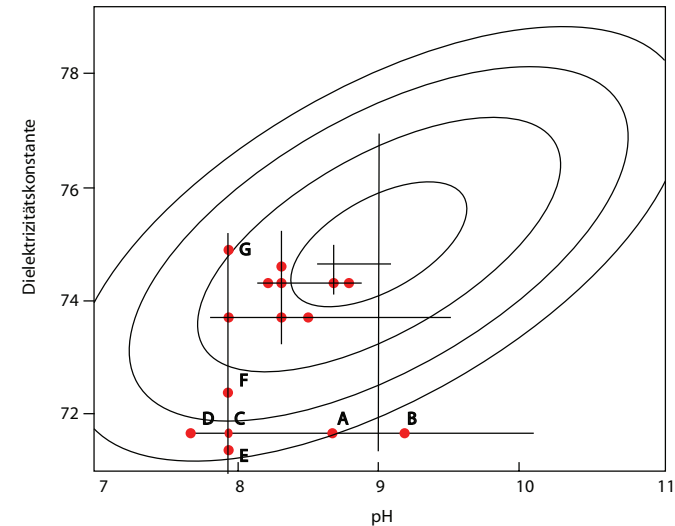
Ein Faktor pro Schritt



Ein Faktor pro Schritt

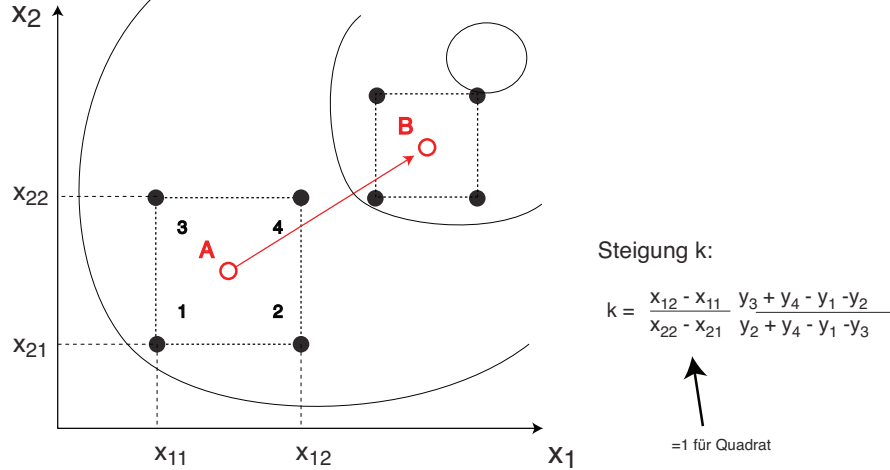


Ein Faktor pro Schritt



Schätzung der ersten Ableitung

Bei der Methode von Box und Wilson werden Versuche nach einem 2^n -faktoriellen Plan durchgeführt und die Richtung der neuen Messungen (nicht aber die Schrittgröße) aufgrund der ersten Ableitung bestimmt.



Simplex

In einem N-dimensionalen Faktorenraum startet man mit N+1 Experimenten, jedes bei einer anderen Kombination der Faktoren. Die N+1 Punkte definieren den Startsimplex.

Aufgrund der Antworten errechnet man aus den Faktorenwerten des Startsimplex die Faktoren des nächsten Experiments. Dazu bestimmt man den Schwerpunkt "centroid" wie folgt: Von den N+1 Faktorenkombinationen wird diejenige mit der schlechtesten Antwort gestrichen. Der Durchschnitt der anderen Faktorenkombinationen definiert den Schwerpunkt.

Für die nächste Messung erhält man die Faktorenkombination durch Spiegelung der Faktorenwerte der schlechtesten Antwort am Schwerpunkt.

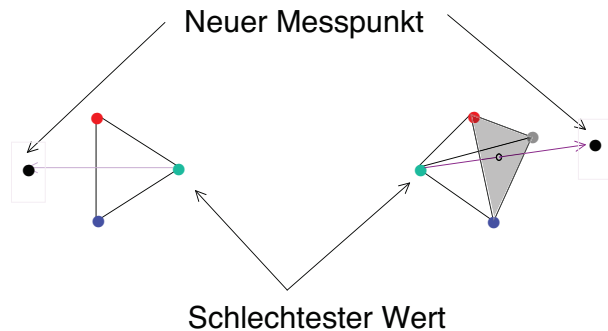
Die weiteren Schritte erfolgen nach dem gleichen Prinzip.

Simplex

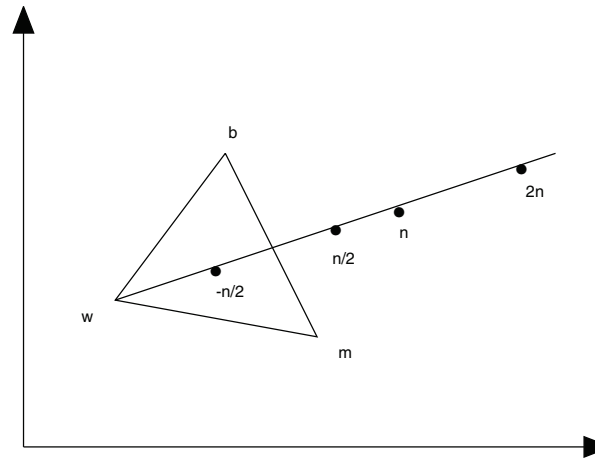
Der Name der Methode "Simplex" bezeichnet den einfachsten konvexen Polyeder eines Raumes gegebener Dimension.

Für 2 Dimensionen: Dreieck

Für 3 Dimensionen: Tetraeder



Modifizierter Simplex

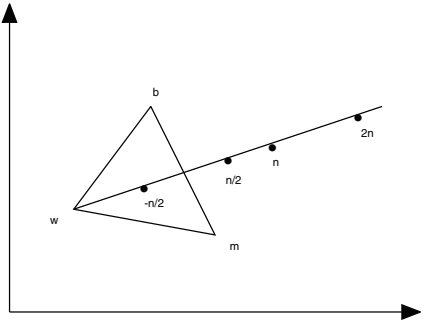


Erste drei Messungen: b: best, m: medium, w: weakest
Neue Messung nach Standardverfahren bei n

J.A. Nelder and R. Mead, A simplex method for function minimization, Computer Journal, 7, 308-313 (1965).

E. Morgan and K.W. Burton, Optimization using the super-modified simplex method, Chemom. Intell. Lab. Systems, 8, 97-107 (1990).

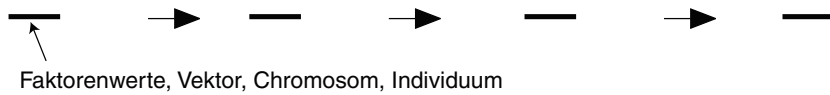
Modifizierter Simplex



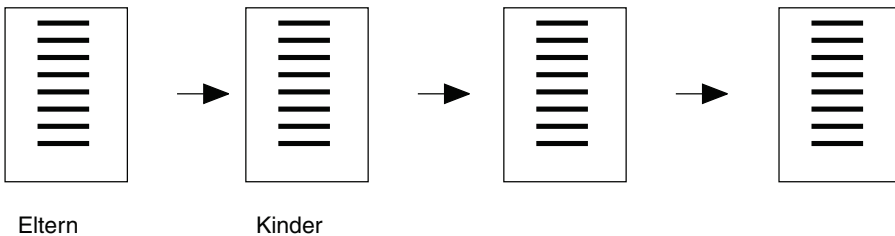
Neuer Simplex
 wenn $m < n < b$ $m \ n \ b$
 wenn $b < n$:
 dann neue Messung bei $2n$
 wenn $b < 2n$ $m \ 2n \ b$
 wenn $2n < b$ $m \ n \ b$
 wenn $w < n < m$ $m \ n/2 \ b$
 wenn $n < w$ $m \ -n/2 \ b$

Einzelversuche und Populationen

Optimierung mit Einzelversuchen



Optimierung mit Populationen



Stochastische Optimierungsmethoden

Mit Einzelversuchen: Simulated Annealing

Mit Populationen: Evolutorische Algorithmen
 Genetische Algorithmen

Vorteil: Lokale Minima können überwunden werden

Nachteil: Nahe beim Optimum nicht effizient --> Hybridmethoden

Zur Erinnerung: Keine Methode kann das Auffinden des globalen Optimums garantieren.

Simulated Annealing

- Bei einer zufällig generierten Faktorenkombination (**a**) wird „gemessen“ (Antwort E_a).
- Die Faktoren werden durch zufällige Störungen verändert.
- Eine neue Faktorenkombination (**n**) wird mit der relativen Wahrscheinlichkeit P akzeptiert:

$$P = 1, \text{ wenn } E_n < E_a$$

$$P = e^{-(E_n - E_a)/c}, \text{ wenn } E_n > E_a$$

c Temperaturparameter, wird während der Optimierung schrittweise verkleinert
- Auswirkung: lokale „Energie“-Minima können überwunden werden. Durch die schrittweise Verkleinerung von c nimmt die Wahrscheinlichkeit ab, dass ein schlechterer Wert akzeptiert wird.

*Annealing: Glühen, Härten, Tempern

Genetische Algorithmen

Optimierung mit Populationen statt Einzelzuständen.

Die Startpopulation wird durch zufällige Faktorenkombinationen erzeugt.

Die nächste Generation wird durch die genetischen Operationen erzeugt:

Mutation
Cross-Over

Die Wahrscheinlichkeit, dass ein Individuum bei der Erzeugung der neuen Population zu den Eltern mit aufgenommen wird, hängt von seiner Qualität (Fitness) ab.

Die Faktoren können binär oder reelle Zahlen sein.

Mutation

Bei der Binärkodierung bedeutet die Mutation, dass an den zufällig ausgewählten Positionen 0 durch 1 und 1 durch 0 ersetzt wird.

Bei reeller Kodierung wird zu den Faktoren bei den zufällig ausgewählten Positionen jeweils eine Zufallszahl addiert. Der Bereich der Zufallszahlen soll der Aufgabe sinnvoll angepasst werden, z. B. -30° bis $+30^\circ$ bei einem dihedralen Winkel während der Optimierung einer Konformation.

Cross-Over

Beim Cross-Over werden gewisse zufällig ausgewählte Faktoren zwischen zwei für die Erzeugung eines Kindes ausgewählten Individuen ausgetauscht.

“Single-point”-Cross-Over:

aaaaalaaa	→	aaaaabbb
bbbbblbbb		bbbbbaaa

Ein Cross-Over-Punkt wird zufällig gewählt.

“Uniform”-Cross-Over

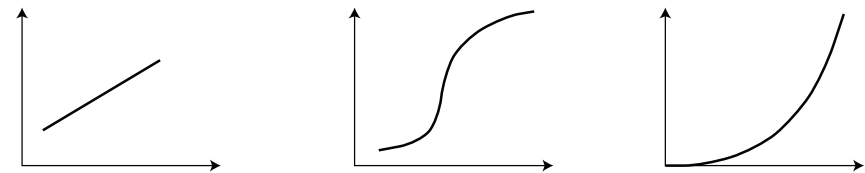
aaaaaaaa	→	abaaabba
^_^_^		
bbbbbbbbb		babbbaab

Mehrere Bereiche werden zufällig gewählt. Das zweite Kind ist das Gegenteil des ersten.

Fitnessfunktion

Das Qualitätsmerkmal wird oft in eine Fitnessfunktion transformiert.

Einige übliche Transformationsfunktionen sind:



Die Fitnessfunktion kann während der Optimierung automatisch angepasst werden.

Selektionsmethoden

Roulette-Rad: Die Wahrscheinlichkeit der Selektion ist der Fitness proportional.

“Linear Ranking”: Die Wahrscheinlichkeit der Selektion hängt von der Position in der Rangliste ab.

“Truncation Selection”: Die besten n Individuen werden mit gleicher Wahrscheinlichkeit ausgewählt.

“Tournament Selection”: Es werden zufällig n Individuen ausgewählt. Das beste wird gewählt.

Sharing

Es ist vorteilhaft, wenn ein Individuum einer Population nahe beim Optimum liegt, aber die anderen Individuen andere Teile der Suchfläche belegen.

Deshalb wird oft eine zusätzliche Mutation eingeschaltet, wenn ein neu erzeugtes Individuum einem anderen zu ähnlich ist.

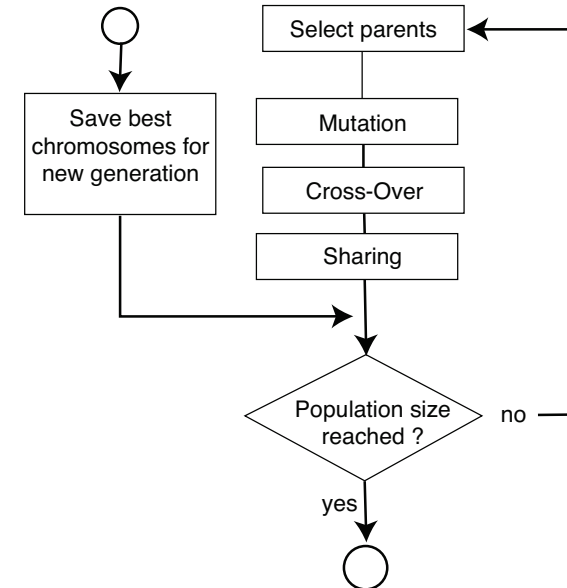
Dadurch verhindert man, dass alle Individuen in die Nähe des gleichen Minimums zu liegen kommen (premature convergence). Diese Strategie nennt man (etwas irreführend) "sharing" (sharing: teilend, gemeinsame Benutzung).

Generation gap, Elitism

Um die besten Individuen nicht zu verlieren, werden die n besten einer Population unverändert in die nächste Generation übernommen.

n wird typischerweise klein gewählt.

Flussschema eines GAs



Auswahl von Wellenlängen, Literatur

(1) Wavelengths selection and optimization of pattern recognition methods using the genetic algorithm.

Smith, B. M.; Gemperline, P.J.

Anal. Chim. Acta **2000**, 423, 167-177.

(2) Genetic algorithms as a tool for wavelength selection in multivariate calibration.

Jouan-Rimbaud, D.; Massart, D.-L.; Leardi, R.; De Noord, O. E.

Anal. Chem. **1995**, 67, 4295-4301.

(3) Genetic algorithm-based wavelength selection for near infrared determination of glucose in biological matrices: Initialization strategies and effects of spectral resolution.

Ding, Q.; Small, G. W.; Arnold, M. A.

Anal. Chem. **1998**, 70, 4472-79.