

6. Experimentieren mit Simulatoren

Wie im vorherigen Kapitel Simulation als Basis oder Hilfsmittel zur Entscheidungsunterstützung

Im vorherigen Kapitel Auswahl einer Konfiguration aus einer Menge von Konfigurationen

Hier nun: Einstellen der Systemparameter, so dass ein bestimmtes Ziel erreicht wird

Unterschied zum bisherigen Vorgehen:

- Zahl und Definitionsbereich der Parameter, die potenziell modifizierbar sind, führt zu sehr vielen (u.U. unendlich vielen) Konfigurationen
- Es können nicht alle Konfigurationen analysiert/simuliert werden
- Finden einer möglichst aussagekräftigen Teilmenge von Konfigurationen zur Herleitung der notwendigen Aussagen

Ziele des Kapitels:

- Einführung in die Methodik des Designs von Experimenten
- Einordnen und Klassifizieren von Faktoren
- Kennen lernen von Faktordesigns
- Kennen lernen von Methoden zur Behandlung von Systemen mit vielen Faktoren
- Überblick über weitere Methoden des Experimentdesigns
- Kennen lernen von Screening Methoden zur Faktorauswahl
- Erkennen von Simulationsspezifika beim Experimentieren
- Einführung in Methoden zur Metamodellbildung
- Einführung in die Varianzanalyse
- Überblick über Response Surfaces

Literatur:

- Law/Kelton Kap. 12.1-12.4.
- D. C. Montgomery.
Design and Analysis of Experiments. Wiley 1997.
- R. H. Myers, D. C. Montgomery
Response Surface Methodology. Wiley 2002.

Kapitelgliederung

- 6.1 Einführung in das Experimentieren
- 6.2 Experimente mit einem Faktor
- 6.3 Experimentdesigns mit wenigen Faktoren
- 6.4 Behandlung von großen Faktorzahlen im Experimentdesign
- 6.5 Weitere Experimentdesigns
- 6.6 Metamodelle
- 6.7 Auswahl relevanter Faktoren und Faktorkombinationen

6.1 Einführung in das Experimentieren

Problemstellung:

Für ein System, welches

- auf unterschiedliche Arten konfiguriert, organisiert, betrieben werden kann (indem die Werte der kontrollierbaren Größen $w_C \in W_C$ festgelegt werden)
- durch die Werte weiterer nicht kontrollierbarer Größen beeinflusst wird (indem Größen $w_U \in W_U$ ihren Wert ändern)

soll das Verhalten bzgl. eines Leistungsmaßes

$$v = g(w_P) = g(f(w_C, w_U)) \text{ beurteilt werden}$$

Simulator ist ein numerisches Modell, d.h.

- g ist nicht explizit darstellbar, sondern
- v kann nur für bestimmte Situationen beobachtet werden

Methodik des Vorgehens:

- Setzen von Situationen (d.h. w_C festlegen)
- Beobachten der Resultate bei unterschiedlicher Realisierung von w_U (Realisierung des Zufalls vorgegeben durch Saat des ZZ-Generators)
- **Experimentieren** mit dem Simulator
- **Design von Experimenten** wird benötigt
(Plan, welche Experimente durchzuführen sind)

Experimentieren ist ein übliches Vorgehen in den Naturwissenschaften:

- Definition von Zielen
- Festlegung von zur untersuchenden Konfigurationen (Experimentplan)
- Durchführung der Experimente
- U.U. Definition neuer Experimente

Bespielanwendung	Eingaben	Ausgaben
Chemische Reaktion	Druck, Temperatur, Katalysator	Menge, Qualität
Anbau von Tomaten	Dünger, Wasser, Saatgut, PH-Wert des Bodens	Menge, Qualität
Simulation eines Fertigungssystems	Anzahl Maschinen, Zuverlässigkeit der Maschinen, Auswahlstrategien der Teile, Puffergrößen	Durchsatz, Auslastung, Verweilzeiten

Unterschiede Simulation – reale Experimente:

- Beliebige Kontrollierbarkeit auch eigentlich unkontrollierbarer Größen
- Einfache Wiederholung von Experimenten
- Beobachtbarkeit aller Größen

Beim Experimentieren werden

- gewisse Komponenten aus W_C fest eingestellt
 - andere Komponenten aus W_C werden systematisch verändert
- veränderliche Komponenten heißen **Faktoren**

Man unterscheidet:

- Qualitative Faktoren
 - können ein Niveau aus einer endlichen Menge von Niveaus annehmen
 - Sind nicht zwingend geordnet (z.B. unterschiedliche Schedulingstrategien, Winter/Sommer)
- Quantitative Faktoren
 - können einen Wert aus einem Kontinuum annehmen (z.B. Regenmenge, Geschwindigkeit)

Größen mit diskretem Wertevorrat können als quantitative oder qualitative Größen behandelt werden

System	Faktoren	Quant	Qual	Kont	Unkont	Resultat
Supermarkt	Ankunftszeit	X			X	Verweilzeit
	Bedienzeit	X		X?		Schlangenlänge
	Anzahl Kassen	X?	X?	X		Kassen-
	Expresskasse		X	X		auslastung
Fertigungs- linie	Anzahl Maschinen	X?	X?	X?	X?	Durchsatz
	Abfertigungsdis.		X	X?	X?	Zeit im System
	Puffergröße	X?	X?	X?	X?	Auslastung
	Maschinengrupp.		X	X?	X?	Gewinn
	Geschw. Förderband	X		X?	X?	
Einkaufs- strategie	Auftragsrate	X			X	Umsatz
	Auftragsgröße	X			X	Lagerkosten
	Lieferzeit	X		X?	X?	Verlorene Aufträge
	Bestellzeiten	X		X		Gewinn
	Max. Lagerbestand	X		X?	X?	

Beobachtungen:

1. Ob ein Faktor kontrollierbar oder unkontrollierbar ist, ist situationsabhängig zu entscheiden
2. Quantitative Faktoren, die geordnet sind, können in qualitative Faktoren transformiert werden, indem z.B. nur der kleinste und größte Wert aus einem plausiblen Intervall betrachtet werden.

Auf Grund von 2. beschäftigen sich Methoden zum Experimentdesign oft mit qualitativen Faktoren

Oftmals werden nur zwei Niveaus (groß – klein) betrachtet

Dabei wird das betrachtete Intervall u.U. in Abhängigkeit von Resultaten früherer Experimente modifiziert
(interessante Größen werden eingeschachtelt)

Unser Modell zum Experimentieren:

$$v = g(t^*, w, s)$$

mit t^* Zeitpunkt der Resultatbetrachtung
 w Vektor der Einflussgrößen
 s Saatwert(e) des ZZ-Generators

Typische Fragestellungen:

- Welche Situationen w sind zu untersuchen?
- Wie viele Replikationen sind für ein w zu beobachten?
- Was kann man bzgl. des Vergleichs der Situationen w und w' aussagen?
- Wie muss man w verändern, um v zu optimieren?

Beantwortung der Fragen durch Methoden des
Experimentdesigns und der **Experimentanalyse**
(Design of experiments (DOE))

6.2 Experimente mit einem Faktor

Situation: Ein Faktor, der für m verschiedene Werte (Niveaus) untersucht werden soll

- x_i ist das i -te Niveau des Faktors
- $\mu_i = f(x_i)$ der zu beobachtende Wert bei Niveau x_i
- Falls $f(\cdot)$ durch ein Simulationsmodell realisiert wird, ist μ_i nur mit statistischen Schwankungen zu beobachten
- Übliches Vorgehen beobachte n Replikationen
 $y_{ij} = f(x_i, s_j)$ für jedes Niveau, wobei die s_j so gewählt sind, dass unabhängige Ströme von Zufallszahlen entstehen

Ziel: Unterscheidung der statistischen Schwankungen von „wahren“ Unterschieden in den Resultaten, insbesondere gilt $\mu_1 = \mu_2 = \dots = \mu_m$?

Vergleich der Mittelwerte mit Hilfe der Methoden aus Kapitel 5

- Für $m=2$ nutze Methoden aus 5.1 zum Vergleich der Systemkonfigurationen
(d.h. bestimme Konfidenzintervall für $\tilde{W} = \tilde{\mu}_1 - \tilde{\mu}_2$)
- Für $m>2$ Methoden aus 5.2 oder 5.3 verwenden
Diese sind aber aufwändig, wenn kleine Konfidenzintervalle notwendig sind

Alternative: Testverfahren

$H_0: \mu_1 = \mu_2 = \dots = \mu_m$ und $H_1: \mu_k \neq \mu_1$ für mindestens ein Paar (k,l)

Teststatistik

$$F_0 = \frac{nm(n-1) \sum_{i=1}^m (\tilde{\mu}_i - \tilde{\mu})^2}{(m-1) \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \tilde{\mu}_i)^2} \quad \text{mit} \quad \tilde{\mu} = \frac{1}{m} \sum_{i=1}^m \tilde{\mu}_i$$

Falls $F_0 > F_{\alpha, m-1, nm-m}$ kann H_0 zum Niveau α verworfen werden

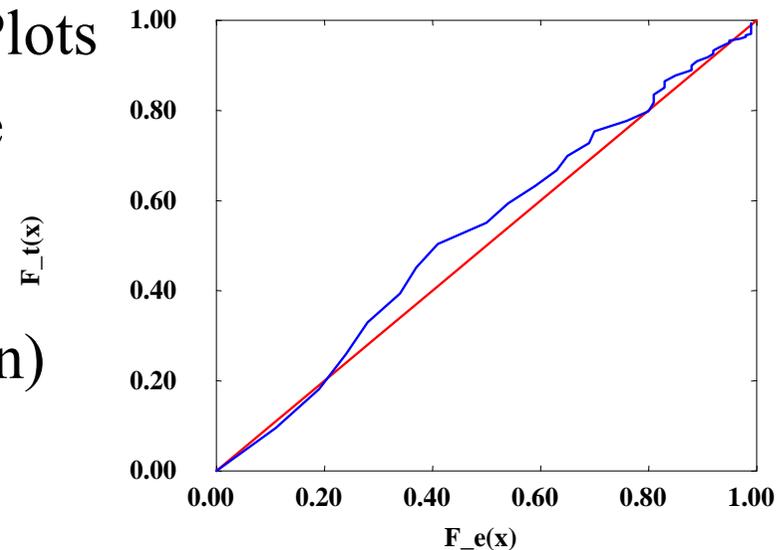
\Rightarrow Niveaus des Faktors haben einen signifikanten Einfluss!

Testverfahren liefern nur dann zuverlässige Resultate, wenn die Annahmen gelten

Kann man dies feststellen?

Zwei Arten von Verfahren:

- Visualisierung der (geordneten) Residuen gegen eine $N(0, \sigma)$ -Verteilung durch P-P oder Q-Q Plots
 - Problematisch kann der kleine Stichprobenumfang sein (kann durch zusätzliche Replikationen behoben werden)
 - Subjektive Entscheidung auf Basis des visuellen Eindrucks
- Alternative: Diverse Testverfahren



6.3 Experimentdesigns mit wenigen Faktoren

- k (≥ 2) Faktoren mit jeweils m_i ($i = 1, \dots, k$) Niveaus
jeweils ein Ausgangsniveau und $m_i - 1$ weitere Niveaus
- **Einfaches Design:**
Simuliere das Modell wobei alle Faktoren auf dem
Ausgangsniveau sind
Ändere pro Experiment genau eine Faktor
 $N = \sum_{i=1, \dots, k} (m_i - 1)$ Experimente sind notwendig
Aber Einfluss der Faktoren ist oft nicht unabhängig
z.B. schnellere CPU + größerer Hauptspeicher in einem PC
- **Vollständiges Design:**
Alle Kombinationen von Faktorniveaus werden untersucht
 $N = \prod_{i=1, \dots, k} m_i$ Experimente sind notwendig
alle Effekte berücksichtigt aber viele Experimente notwendig

Zur Reduktion der Zahl der Experimente:

- Zahl der Niveaus reduzieren
 - oft werden nur zwei Niveaus erfasst, die als -1 und +1 bezeichnet werden
 - zu Beginn werden (nicht zu große) Intervalle für Faktorwerte geschätzt und das Minimum und Maximum als Niveaus verwendet
 - in iterativen Experimenten können sich die Intervalle ändern (z.B. zur Optimierung)
- Zahl der Faktoren reduzieren
 - Zuerst untersuchen, ob ein Faktor überhaupt einen Einfluss auf das Modellverhalten hat (nicht einfach, da Abhängigkeiten existieren)
- Teilentwürfe verwenden
 - Kompromiss zwischen einfachem und vollständigem Design

Faktordesigns für k Faktoren:

Aufzählen aller Vektoren $(-1, \dots, -1), (+1, -1, \dots, -1), \dots, (+1, \dots, +1)$

Änderung der ersten Stelle zuerst

$\Rightarrow 2^k$ Experimente (von 1 bis 2^k nummeriert)

R_i sei das Resultat des i-ten Experiments

Experimente für den Fall $k=3$:

Exp.	F1	F2	F3	Exp.	F1	F2	F3
1	-1	-1	-1	5	-1	-1	+1
2	+1	-1	-1	6	+1	-1	+1
3	-1	+1	-1	7	-1	+1	+1
4	+1	+1	-1	8	+1	+1	+1

e_i sei der Effekt von Faktor i

Diese Effekte für $i=1,2,3$ werden auch als Haupteffekte bezeichnet

- Haupteffekte beschreiben die durchschnittliche Änderung des Resultats, wenn Faktor i von -1 nach $+1$ wechselt und alle anderen Faktoren unverändert bleiben

Es gilt:

$$e_1 = \frac{(R_2 - R_1) + (R_4 - R_3) + (R_6 - R_5) + (R_8 - R_7)}{4}$$

$$e_2 = \frac{(R_3 - R_1) + (R_4 - R_2) + (R_7 - R_5) + (R_8 - R_6)}{4}$$

$$e_3 = \frac{(R_5 - R_1) + (R_6 - R_2) + (R_7 - R_3) + (R_8 - R_4)}{4}$$

Neben den Haupteffekten (= Einfluss eines einzelnen) treten Nebeneffekte durch die Interaktion von zwei oder mehr Faktoren auf

Messung der Interaktion zwischen zwei Faktoren i und j:

- Hälfte der Differenz zwischen dem mittleren Effekt von Faktor i, wenn Faktor j das Niveau +1 hat (alle anderen Faktoren konstant) und wenn Faktor j das Niveau -1 hat (alle anderen Faktoren konstant)

$$e_{12} = \frac{1}{2} \left(\frac{(R_4 - R_3) + (R_8 - R_7)}{2} - \frac{(R_2 - R_1) + (R_6 - R_5)}{2} \right)$$

$$e_{13} = \frac{1}{2} \left(\frac{(R_6 - R_5) + (R_8 - R_7)}{2} - \frac{(R_2 - R_1) + (R_4 - R_3)}{2} \right)$$

$$e_{23} = \frac{1}{2} \left(\frac{(R_7 - R_5) + (R_8 - R_6)}{2} - \frac{(R_3 - R_1) + (R_4 - R_2)}{2} \right)$$

Interpretation von Interaktionen höherer Ordnung ist schwieriger, kann aber mathematisch analog definiert werden!

Messung der Interaktion zwischen drei Faktoren h, i und j:

- Hälfte der Differenz zwischen der durchschnittlichen Interaktion zwischen h und i, wenn j Niveau +1 hat und der Interaktion zwischen h und i, wenn j Niveau -1 hat
(wie vorher werden weitere Faktoren konstant gehalten)

Bei drei Faktoren existiert nur eine Interaktion der Ordnung 3:

$$e_{123} = \frac{1}{2} \left(\frac{(R_8 - R_7) + (R_5 - R_6)}{2} - \frac{(R_4 - R_3) + (R_1 - R_2)}{2} \right)$$

- Konzept auf weitere Interaktionen erweiterbar
- Kompakte Darstellung der Interaktionen in der Designmatrix

Betrachtung des Falls $k=3$

Designmatrix (wird oft mit X bezeichnet):

Design Punkt	Effekte							Response
	e_1	e_2	e_3	e_{12}	e_{13}	e_{23}	e_{123}	
1	-1	-1	-1	+1	+1	+1	-1	R_1
2	+1	-1	-1	-1	-1	+1	+1	R_2
3	-1	+1	-1	-1	+1	-1	+1	R_3
4	+1	+1	-1	+1	-1	-1	-1	R_4
5	-1	-1	+1	+1	-1	-1	+1	R_5
6	+1	-1	+1	-1	+1	-1	-1	R_6
7	-1	+1	+1	-1	-1	+1	-1	R_7
8	+1	+1	+1	+1	+1	+1	+1	R_8

Es bleiben einige Fragen:

1. Wie geht man mit stochastischen Schwankungen in den Resultaten um?
2. Haben alle Faktoren und Effekte wirklich einen Einfluss?
3. Wie behandelt man große Faktorzahlen?
4. Welche Faktoren sind besonders wichtig?

3. und 4. werden in den Abschnitten 6.4 und 6.7 behandelt

Im Folgenden kurze Betrachtung von 1. und 2.

Übliches Vorgehen bei stochastischen Schwankungen:

- Mehrere Replikationen pro Faktorkombination
- Resultate als Mittelwerte der Replikationen
- Schätzung der Varianz zur Beurteilung der Qualität der Resultate

Vorgehen zur Berechnung von Konfidenzintervallen nach dem bekannten Schema:

- Es werden n Replikationen für jede Faktorkombination durchgeführt
- Sei R_i^z das Resultat der i -ten Faktorkombination in der z -ten Replikation
- e_j^z ist der ermittelte Haupteffekt für Faktor j in der z -ten Replikation (analog können Nebeneffekte definiert werden)

Damit gilt:
$$\tilde{e}_i = \frac{1}{n} \sum_{z=1}^n e_i^z \quad \text{und} \quad \tilde{S}^2(\tilde{e}_i) = \frac{1}{n-1} \sum_{z=1}^n (e_i^z - \tilde{e}_i)^2$$

Mit dem Konfidenzintervall
$$\tilde{e}_i \pm t_{n-1, 1-\alpha/2} \cdot \sqrt{\frac{\tilde{S}^2(\tilde{e}_i)}{n}}$$

- Vorgehen erlaubt die Berechnung von Konfidenzintervallen für alle Haupt- und Nebeneffekte
- Damit kann u.a. geprüft werden, ob die Effekte signifikant sind (d.h. ob 0 im Konfidenzintervall enthalten ist oder nicht)

- Aber Vorsicht:

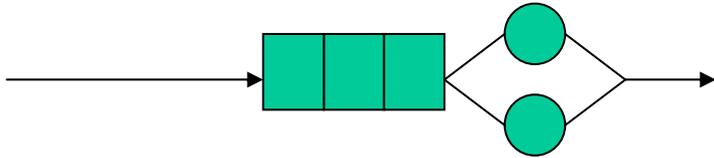
Die einzelnen Konfidenzintervalle sind nicht unabhängig!

Unabhängige Konfidenzintervalle erfordern:

- Nutzung von unabhängigen Replikationen pro Effekt (d.h. $n \cdot 2^{2k}$ Replikationen) oder
- Nutzung der Bonferroni-Ungleichung:

$$1 - \alpha \geq 1 - \sum \alpha_i$$
wobei α_i die Signifikanzwahrscheinlichkeiten der einzelnen Effekte sind und α die Signifikanzwahrscheinlichkeit des Gesamtergebnisses

Ein einfaches Beispiel:



- Ankunftsstrom Poisson mit Rate 1.333
 - Bedienzeiten exponentiell verteilt mit Mittelwert 1.0
 - b Pufferplätze
 - s Bediener
-
- Zielfunktion $f(b,s) = \text{Durchsatz} / (1 + 0.15b + 0.25s)$
 - Niveaus $b \in \{2,3\}$ und $s \in \{1,2\}$
 - Analyse per Simulation
 - jeweils 5 unabhängige Replikationen
 - Länge einer Replikation 1100 Zeiteinheiten, die ersten 100 Zeiteinheiten werden keine Daten erhoben

Bestimmung der Resultate

	Niveaus			
Repl.	(-1,-1)	(+1,-1)	(-1,+1)	(+1,+1)
1	0.483	0.515	0.530	0.574
2	0.501	0.471	0.512	0.588
3	0.482	0.493	0.555	0.581
4	0.488	0.503	0.552	0.556
5	0.488	0.494	0.543	0.584
\tilde{R}	0.488	0.495	0.538	0.577
\tilde{S}	0.00778	0.0164	0.0176	0.0126
90% Konf.In.	± 0.00742	± 0.0156	± 0.0168	± 0.0121

Bestimmung der Haupt- und Nebeneffekte

Repl.	e_1	e_2	e_{12}
1	0.0383	0.0530	0.0129
2	0.0226	0.0643	0.0742
3	0.0185	0.0806	0.0385
4	0.0094	0.0587	0.0189
5	0.0233	0.0719	0.0419
\tilde{e}_i	0.0224	0.0657	0.0373
$\tilde{S}(e_i)$	0.0104	0.0109	0.0241
90% Konf.In.	± 0.00996	± 0.0104	± 0.0229

Ergebnisse zeigen,
dass beide Haupt- und der
Nebeneffekt signifikant
sind
(d.h. 0 ist in keinem
Konfidenzintervall
enthalten)

6.4 Behandlung großer Faktorzahlen

- Die Anzahl der auszuführenden Experimente wächst schnell mit wachsender Faktor- und Replikationszahl
- Realistische Modelle können 10 bis 1000 Faktoren umfassen, wobei jedes Experiment für sich schon mehrere Minuten benötigt
- Experimente sind kaum noch durchführbar (auch bei Nutzung paralleler Hardware)

Was kann man tun, um den Aufwand zu reduzieren?

1. Unwichtige Faktoren aussortieren
 2. Weniger als $r \cdot 2^k$ Experimente bei k Faktoren durchführen
1. Betrachten wir in 6.7, 2. wird hier untersucht

Was bedeutet es, weniger als $r \cdot 2^k$ Experimente durchzuführen?

- Anzahl Replikationen pro Experiment sollte nicht verringert werden, da dies zu ungenauen Ergebnissen führt
- Nicht alle möglichen Faktorkombinationen analysieren

Übliches Vorgehen:

Entwicklung von 2^{k-p} Designs, d.h.

- die Zahl der Experimente wird um den Faktor $1/2^p$ reduziert
- es werden gewisse Faktorkombinationen weggelassen

Die Entwicklung von 2^{k-p} Designs wird in der Literatur über Experimentplanung breit behandelt

Hier wird nur eine kurze Einführung gegeben

Weitere Details z.B. in den Büchern von Montgomery.

Um die vorgestellten Berechnungen durchzuführen, möchte man in der Regel ein orthogonales Design haben, d.h.

➤ die Spalten der Designmatrix X sind paarweise orthogonal

Weiterhin fordern wird

➤ die Summe der Elemente in jeder Spalte ist 0

Einfache Konstruktionsvorschrift für ein 2^{k-p} Design

➤ Generiere ein $2^{(k-p)}$ Design

➤ Ordne die fehlenden Haupteffekte den Spalten $k-p+1$ bis k zu

- damit lassen sich unterschiedliche Designs einfach realisieren
- die Spalten der Designmatrix bleiben paarweise orthogonal
- verschiedene Effekte werden identischen Spalten zugeordnet (dies ist beim vollständigen Design nicht der Fall!)

Beispiel eines 2^{4-1} Designs

Design Punkt	Effekte							Response
	e_1	e_2	e_3	e_{12}	e_{13}	e_{23}	$e_{123} \rightarrow e_4$	
1	-1	-1	-1	+1	+1	+1	-1	R_1
2	+1	-1	-1	-1	-1	+1	+1	R_2
3	-1	+1	-1	-1	+1	-1	+1	R_3
4	+1	+1	-1	+1	-1	-1	-1	R_4
5	-1	-1	+1	+1	-1	-1	+1	R_5
6	+1	-1	+1	-1	+1	-1	-1	R_6
7	-1	+1	+1	-1	-1	+1	-1	R_7
8	+1	+1	+1	+1	+1	+1	+1	R_8

Aus dem Design folgt:

$$e_4 = e_{123} = \frac{1}{2} \left(\frac{(R_8 - R_7) + (R_5 - R_6)}{2} - \frac{(R_4 - R_3) + (R_1 - R_2)}{2} \right)$$

Der Effekt beschreibt also eigentlich $e_4 + e_{123}$

In ähnlicher Form gilt

$$e_{12} = e_{34} = \frac{1}{2} \left(\frac{(R_1 - R_2) + (R_5 - R_6)}{2} - \frac{(R_3 - R_4) + (R_7 - R_8)}{2} \right)$$

Der Effekt beschreibt also eigentlich $e_{12} + e_{34}$

In ähnlicher Form gelten folgende weitere Identitäten:

$$\begin{array}{llll} e_1 = e_2 e_3 e_4 & e_3 = e_1 e_2 e_4 & e_1 e_2 = e_3 e_4 & e_2 e_3 = e_1 e_4 \\ e_2 = e_1 e_3 e_4 & e_4 = e_1 e_2 e_3 & e_1 e_3 = e_2 e_4 & 1 = e_1 e_2 e_3 e_4 \end{array}$$

- Man spricht in diesen Fällen davon, dass die entsprechenden Effekte vermengt (Engl. confounded) sind
 - Dies bedeutet, dass aus den Schätzern der Effekte nicht auf den einzelnen Effekt geschlossen werden kann, sondern nur auf kombinierte Effekte

Die Auflösung eines Designs gibt an, welche Effekte nicht vermengt werden

So sind in einem Design der Ordnung IV keine Effekte der Ordnung p und q mit $p+q < 4$ vermengt

In unserem Beispiel sind

- Haupteffekte (Ordnung 1) nicht untereinander und auch nicht mit Nebeneffekten der Ordnung 2 vermengt,
- Nebeneffekte der Ordnung 2 sind dagegen vermengt

damit hat das Design die Ordnung IV

Man schreibt die Auflösung als römische Zahl

In unserem Fall liegt also ein 2_{IV}^{4-1} Design vor

Anhaltspunkte zur Konstruktion von Designs:

- Der Einheitsvektor sollte mit Effekten höherer Ordnung vermengt werden
- Haupteffekte sollten nur mit Effekten höherer Ordnung vermengt werden
 - In beiden Fällen ist die implizite Annahme, dass die Effekte höherer Ordnung vernachlässigbar sind
- Falls angenommen werden kann, dass bestimmte Nebeneffekte bedeutsam sind, so sollten dies nur mit Effekten höherer Ordnung vermengt werden
- Ein Design höherer Ordnung ist in der Regel besser als ein Design niedrigerer Ordnung, erfordert aber auch mehr Experimente

Beispiele für Designs:

k	Auflösung		
	III	IV	V
3	2^{3-1}_{III} $3 \leftrightarrow 12$		
4		2^{4-1}_{IV} $4 \leftrightarrow 123$	
5	2^{5-2}_{III} $4 \leftrightarrow 12$ $5 \leftrightarrow 13$		2^{5-1}_V $5 \leftrightarrow 1234$
6	2^{6-3}_{III} $4 \leftrightarrow 12$ $5 \leftrightarrow 13$ $6 \leftrightarrow 23$	2^{6-2}_{IV} $5 \leftrightarrow 123$ $6 \leftrightarrow 234$	
8		2^{8-4}_{IV} $5 \leftrightarrow 234$ $6 \leftrightarrow 134$ $7 \leftrightarrow 123$ $8 \leftrightarrow 124$	2^{8-2}_V $7 \leftrightarrow 1234$ $8 \leftrightarrow 1256$

6.5 Weitere Experimentdesigns

- Bisher vorgestellte Designs untersuchen jeweils Faktoren, die „Werte“ -1 und +1 annehmen
- Bei kontinuierlichen Parametern bedeutet dies, dass jeweils der kleinste und größte Wert eines Intervalls gewählt wird
- Es kann gezeigt werden, dass für die Anpassung von linearen Regressionsmodellen, diese Auswahl optimal ist (zu Modellen mit der geringsten Varianz in den Parameterschätzern)
- Trotzdem haben diese Designs auch Nachteile, wenn man mit kontinuierlichen Parametern arbeitet, da
 - Niveaus für Faktoren a priori definiert werden müssen
 - mehr als zwei Niveaus durch unabhängige binäre Variablen kodiert werden

Zufälliges Design:

- Intervalle für die einzelnen Faktoren werden auf $[0,1]$ normiert
- Bei k Faktoren und n Experimenten werden
 - n Zufallszahlen aus $[0,1]^k$ gezogen
 - Verteilung entweder
 - unabhängig und gleichverteilt in allen Dimensionen oder
 - bzgl. vorgegebener Verteilungen, falls Parameterverteilungen bekannt sind
- An jedem Designpunkt wird
 - entweder eine feste Anzahl von Replikationen durchgeführt oder
 - so lange simuliert, bis eine vorgegebene Varianz erreicht wird

Nachteil des zufälligen Designs bei hoher Dimension:

Große Bereiche des Parameterraums werden mit relativ hoher Wahrscheinlichkeit nicht untersucht!

Alternative: Geschichtete Stichproben

- Unterteilung des $[0,1]^k$ in n (ungefähr) gleich große Unterräume $[l_i, u_i]^k$
- Zufällige Auswahl je eines Punktes aus jedem Unterraum
- Verschiedene Varianten existieren
 - Gleichverteilte unabhängige Auswahl in jeder Dimension
 - Auswahl bzgl. vorgegebener Verteilung
 - Unterschiedliche Anzahl Designpunkte in den Unterräumen
 - ...

Latin Hypercube Design:

Eine gute Verteilung der Designpunkte im Raum soll erreicht werden

Vorgehen am Beispiel von zwei Dimensionen und n Experimenten

- Unterteile jede Dimension in n gleich große Intervalle $[0, 1/n), [1/n, 2/n), \dots, [(n-1)/n, 1]$
- Es entsteht ein Gitter mit n^2 Zellen
- Verteile die ersten n Buchstaben so auf die Zellen, dass jeder Buchstabe in jeder Zeile und Spalte einmal vorkommt

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

- Wähle einen Buchstaben
- Bestimme zufällig je einen Designpunkt in jeder zum Buchstaben gehörigen Zelle

Verallgemeinerung auf n Dimensionen:

- Unterteile jede Dimension in k Intervalle gleicher Größe
- Jede Zelle ist durch k ganzzahlige Werte (i_1, \dots, i_k) spezifiziert die Projektion einer Zelle auf Dimension p liefert i_p
- Wähle für das Design n Zellen, so dass für jede Dimension p die Projektionen die Projektion der ausgewählten Zellen die Werte $1, \dots, n$ liefert
- Wähle aus jedem Intervall zufällig einen Punkte

Latin Hypercube Designs können von sehr unterschiedlicher Qualität sein:

			*
		*	
	*		
*			

Schlechte
Verteilung

*			
		*	
			*
	*		

Bessere
Verteilung

6.6 Metamodelle

- Ziel der Metamodellbildung ist die Herstellung eines einfachen Zusammenhangs zwischen den Faktorwerten und der abhängigen Variablen (Response)
- Üblicherweise verwendet werden lineare Regressionsmodelle (Response Surface Models)

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \epsilon$$

- y ist die abhängige Variable
- x_i sind die Faktorwerte
- β_i sind die Regressionskoeffizienten
- ϵ ist der Fehler

Annahme: normalverteilt mit Erwartungswert 0, Varianz σ

Durchführung von n Experimenten liefert n Gleichungen:
 (in der Simulation sind y_i oft Mittelwerte aus mehreren Experimenten)

$$y_i = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_k x_{i,k} + \epsilon_i = \beta_0 + \sum_{j=1}^k \beta_j x_{i,j} + \epsilon_i$$

Ziel: Wähle β_0, \dots, β_k so, dass

$$L = \sum_{i=1}^n (\epsilon_i)^2 = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{i,j} \right)^2 \quad \text{minimal wird}$$

D.h. β_j so wählen, dass partielle Ableitungen 0 werden

$$\left. \frac{dL}{d\beta_0} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left(y_i - b_0 - \sum_{j=1}^k b_j x_{i,j} \right) = 0$$

$$\left. \frac{dL}{d\beta_j} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left(y_i - b_0 - \sum_{j=1}^k b_j x_{i,j} \right) x_{i,j} = 0$$

Einfachere Darstellung in Vektor-Matrix-Schreibweise:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,k} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,k} \end{pmatrix},$$
$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \text{und } \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_k \end{pmatrix}$$

Es gilt:

$$L = \sum_{i=1}^n \epsilon_i = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Etwas anders aufgeschrieben ergibt sich:

$$\begin{aligned} L &= \mathbf{y}^T \mathbf{y} - \beta^T \mathbf{X}^T \mathbf{y} - \mathbf{y} \mathbf{X} \beta + \beta^T \mathbf{X}^T \mathbf{X} \beta \\ &= \mathbf{y}^T \mathbf{y} - 2\beta^T \mathbf{X}^T \mathbf{y} + \beta^T \mathbf{X}^T \mathbf{X} \beta \end{aligned}$$

Mit den partiellen Ableitungen:

$$\left. \frac{dL}{d\beta} \right|_{\mathbf{b}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{0} \Rightarrow \mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Die Elemente in \mathbf{b} sind die “optimalen” Regressionskoeffizienten, die den quadratischen Fehler minimieren

Berechnung von \mathbf{b} erfordert:

- Aufstellen der Matrizen
- Bildung der inversen Matrix (Dimension $k+1 \times k+1$)
- Berechnung Matrix-Matrix- und Matrix-Vektor-Produkt

Aufbau der Matrizen und Vektoren:

$$\mathbf{X}^T \mathbf{X} = \begin{pmatrix} n & \sum_{i=1}^n x_{i,1} & \sum_{i=1}^n x_{i,2} & \cdots & \sum_{i=1}^n x_{i,k} \\ \sum_{i=1}^n x_{i,1} & \sum_{i=1}^n x_{i,1}^2 & \sum_{i=1}^n x_{i,1}x_{i,2} & \cdots & \sum_{i=1}^n x_{i,1}x_{i,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^n x_{i,k} & \sum_{i=1}^n x_{i,k}x_{i,1} & \sum_{i=1}^n x_{i,k}x_{i,2} & \cdots & \sum_{i=1}^n x_{i,k}x_{i,k}^2 \end{pmatrix}$$

$$\mathbf{X}^T \mathbf{y} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i x_{i,1} \\ \vdots \\ \sum_{i=1}^n y_i x_{i,k} \end{pmatrix}$$

Es gilt dann

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} \quad \text{und} \quad \mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

Ein Beispiel:

Untersuchung von 7 Programmen bzgl. des Zusammenhangs zwischen CPU-Zeit (abhängige Variable y) und I/O-Operationen (x_1) sowie Speicherbedarf (x_2):

CPU-Zeit	I/O-Operationen	Speicherbedarf
y_i	$x_{1,i}$	$x_{2,i}$
2	14	70
3	16	75
7	27	144
9	42	190
10	39	210
13	50	235
20	83	400

$$\mathbf{X} = \begin{pmatrix} 1 & 14 & 70 \\ 1 & 16 & 75 \\ 1 & 27 & 144 \\ 1 & 42 & 190 \\ 1 & 39 & 210 \\ 1 & 50 & 235 \\ 1 & 83 & 400 \end{pmatrix} \quad \mathbf{X}^T \mathbf{X} = \begin{pmatrix} 7 & 271 & 1324 \\ 271 & 13855 & 67188 \\ 1324 & 67188 & 326686 \end{pmatrix}$$

$$(\mathbf{X}^T \mathbf{X})^{-1} = \begin{pmatrix} 0.6297 & 0.0223 & -0.0071 \\ 0.0223 & 0.0280 & -0.0058 \\ -0.0071 & -0.0058 & 0.0012 \end{pmatrix}$$

$$\mathbf{X}^T \mathbf{y} = \begin{pmatrix} 66 \\ 3375 \\ 16388 \end{pmatrix} \quad \longrightarrow$$

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = (-0.1614, 0.1182, 0.0265)^T$$

Fehlerberechnung und Qualität des Modells

$$\begin{aligned} E(\mathbf{b}) &= E\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{y}\right) \\ &= E\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T (\mathbf{X}\beta + \epsilon)\right) \\ &= E\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X}\beta + \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \epsilon\right) = \beta \end{aligned}$$

$$Cov(\mathbf{b}) = E((\mathbf{b} - E(\mathbf{b}))(\mathbf{b} - E(\mathbf{b}))^T) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

Die Kovarianzmatrix ist eine symmetrische Matrix der Dimension $k+1 \times k+1$, die an Position (i,j) die Kovarianz zwischen b_i und b_j und an Position i,i die Varianz von b_i beinhaltet

Schätzer für die Varianz

$$\hat{\sigma}^2 = \mathbf{e}^T \mathbf{e} / (n - k - 1)$$

Quadrierter Fehler

$$SS_E = \mathbf{e}^T \mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b})$$

Wichtige Fragestellung:

Welche Koeffizienten beeinflussen das Modellverhalten überhaupt?

Auf Ebene der Regression lautet diese Frage:

Welche β_i unterscheiden sich signifikant von 0?

⇒ Beantwortung mit Hilfe von Testverfahren!

1) Signifikanz der Regression

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \beta_j \neq 0 \text{ für mindestens ein } j \in \{1, \dots, k\}$$

Summe der quadrierten Abweichungen

$$SS_T = \sum_{i=1}^n (y_i - \bar{y})^2 \text{ mit } \bar{y} = 1/n \sum_{i=1}^n y_i$$

$$\text{Annahme } SS_T = SS_R + SS_E$$

Wie groß ist der Anteil SS_R , den die Regression erklärt?

Teststatistik:
$$F_0 = \frac{SS_R/k}{SS_E/(n - k - 1)} = \frac{MS_R}{MS_E}$$

H_0 wird verworfen, falls $F_0 > F_{\alpha, k, n-k-1}$

In unserem Beispiel gilt:

- $SS_E = 5.300$
- $SS_R = 200.41$

}
$$F_0 = 100.205/1.325 = 75.63$$

$$> F_{0.9, 2, 4} = 4.32$$

Die Regression ist signifikant!

2) Signifikanz der einzelnen Regressionsparameter

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0 \text{ für ein festes } j \in \{0, \dots, k\}$$

Falls H_0 gilt, kann x_j aus dem Modell entfernt werden!

$$\text{Teststatistik: } t_0 = \frac{b_j}{\sqrt{\hat{\sigma}^2 \mathbf{C}(j,j)}} \text{ mit } \mathbf{C} = (\mathbf{X}^T \mathbf{X})^{-1}$$

H_0 wird verworfen, falls $|t_0| > t_{1-\alpha/2, n-k-1}$

Test ist nur partiell (und deshalb nur bedingt aussagekräftig), da die Regressionsvariablen miteinander korreliert sind!

Für unser Beispiel gilt:

$$\hat{\sigma}^2 = SS_E/4 = 1.325$$

$$\text{Für } j = 0: \quad t_0 = -0.1614/\sqrt{1.325 \cdot 0.697} = -0.168$$

$$\text{Für } j = 1: \quad t_0 = 0.1182/\sqrt{1.325 \cdot 0.0280} = 0.614$$

$$\text{Für } j = 1: \quad t_0 = 0.0265/\sqrt{1.325 \cdot 0.0012} = 0.665$$

$t_{0.95,4} = 2.132 \Rightarrow$ kein Parameter für sich signifikant

Aber beide Parameter zusammen erklären die abhängige Variable
recht gut

\Rightarrow Einzelne Parameter weglassen (hier z.B. x_1) und eine neue
Regression bilden

Bisher haben wir nur lineare Terme betrachtet, was passiert, wenn gemeinsame Abhängigkeiten z.B. $x_i x_j$ einbezogen werden sollen?

Allgemeine Form mit Nebeneffekten der Ordnung 2:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \sum_{j=i}^k \beta_{i,j} x_i x_j$$

Umformulierung durch Einführung von $(k^2-k)/2$ neuen Variablen
(insgesamt also $k(k+1)/2$ Variablen)

Beispiel $k=2$:

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{1,1} x_1^2 + \beta_{1,2} x_1 x_2 + \beta_{2,2} x_2^2 + \epsilon \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon \end{aligned}$$

⇒ Ansatz wie erläutert anwendbar für beliebige Nebeneffekte

Abhängigkeiten zwischen Faktoren und Ergebnis ist oft nicht linear
⇒ Transformation von Variablen und/oder Ergebnis

Beispiele

Ersetze jeweils y durch z wobei

- $z = y^{1/2}$
- $z = \ln y$
- $z = e^y$
- $z = 1/y$
- ...

Ersetze jeweils x_i durch w_i wobei

- $w_i = (x_i)^{1/2}$
- $w_i = \ln x_i$
- $w_i = \exp(x_i)$
- $w_i = 1/x_i$
- ...

Linear Regression anwendbar, trotz nichtlinearer Zusammenhänge
Systematisches Vorgehen zur Auswahl der besten Transformation fehlt

6.7 Auswahl relevanter Faktoren und Faktorkombinationen

Situation in vielen praktischen Experimenten

- Es gibt eine Vielzahl von Faktoren (oft mehr als hundert),
 - deren Bedeutung vor Beginn der Experimente unklar ist,
 - die teilweise identische Effekte beschreiben,so dass nur ein kleiner Teil der Faktoren ausreichen würde, um die gewünschten Effekte zu beschreiben

Wie können relevante Faktoren von nicht relevanten unterschieden werden?

- Vollständige oder auch 2^{k-p} Designs können aus Komplexitätsgründen nicht angewendet werden
- Selbst einfache Designs sind aufwändig und erlauben keine Aussagen über Faktoren mit identischen Auswirkungen

Vorauswahl von relevanten Faktoren durch Faktor-Screening
zwei Klassen von Ansätzen existieren:

1. Supersaturierte Designs

Anzahl Faktoren ist größer als die Anzahl der Designpunkte

⇒ nur Kombinationen von Faktoren können untersucht werden

Auf Basis der Experimente werden Aussagen über das
Modellverhalten getroffen

2. Gruppenweises Screening

Faktoren werden in Gruppen zusammengefasst und es wird
versucht Gruppen von unwichtigen Faktoren zu ermitteln

⇒ Gruppen von Faktoren werden aussortiert

Neue Experimente mit weniger Faktoren werden durchgeführt,
um so das Modellverhalten zu analysieren