a) Modell mit einfachen Effekten (Effektdarstellung)

$$Y = (e \;\; Z_1^e(C) \ldots Z_{K_1-1}^e(C) Z_1^e(D) \ldots Z_{K_2-1}^e(D)) \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_{K_1-1} \\ \gamma_1 \\ \vdots \\ \gamma_{K_2-1} \end{pmatrix} \tag{4.6}$$

Test auf Effekt von C: $\quad H_0 : \tau_1 = \ldots = \tau_{K_1-1} = 0$

Test auf Effekt von D: $\quad H_0 : \gamma_1 = \ldots = \gamma_{K_2-1} = 0$

Interpretation:
τ_k, γ_l Abweichung vom Gesamtmittel der Kategorien

Beispiel: Designmatrix

Modell mit einem zweistufigen und einem dreistufigen Faktor und jeweils zwei Beobachtungen pro Faktorkombination

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

Beispiel: 2 kategoriale Einflussgrößen

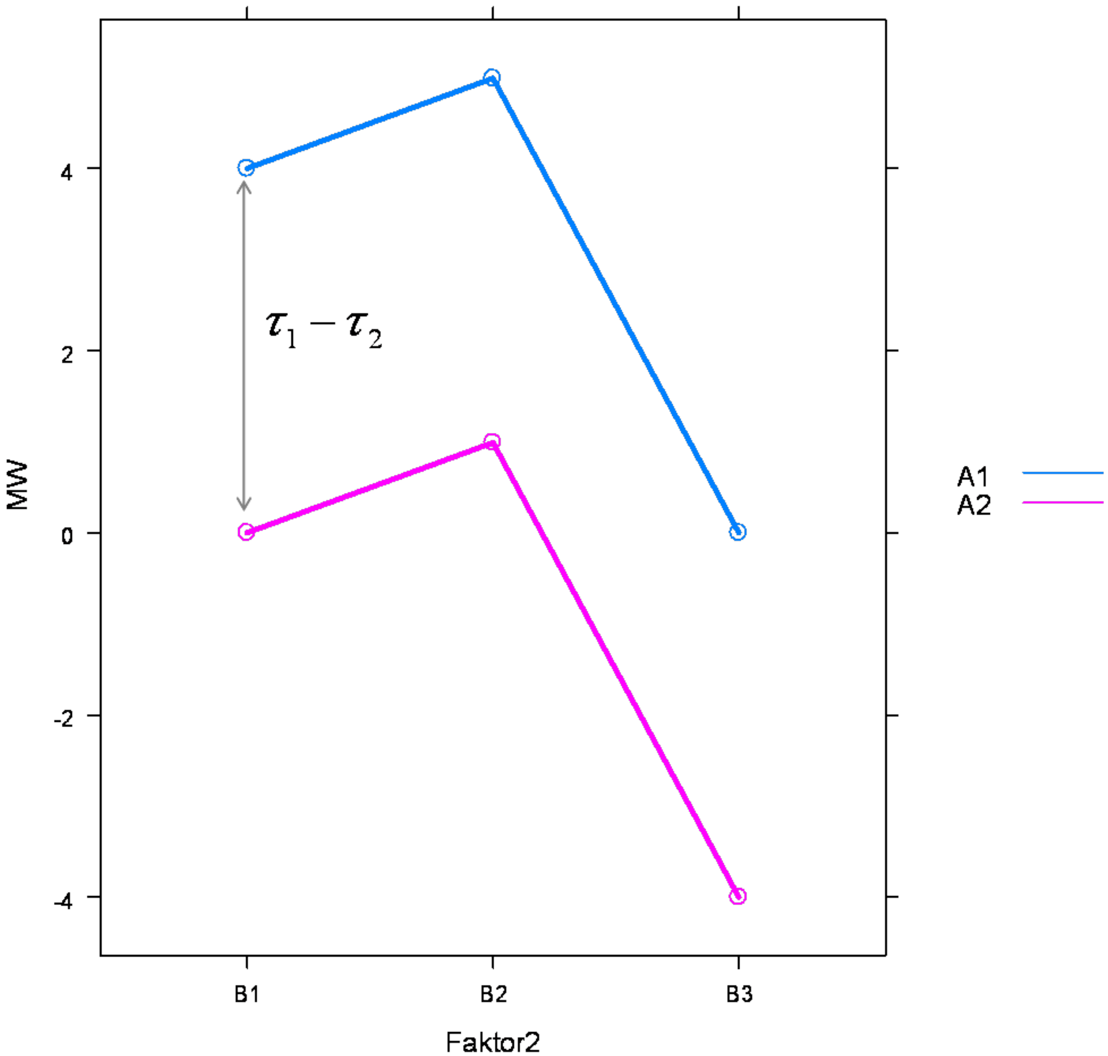
$\mu = 1$

Faktor A (2-stufig):	$\tau_1 = 2$	$\Rightarrow \tau_2 = -2$
Faktor B (3-stufig):	$\gamma_1 = 1, \gamma_2 = 2$	$\Rightarrow \gamma_3 = -3$

Berechnung der Mittelwerte:

Faktor A	Faktor B	MW
1	1	$1 + 2 + 1 = 4$
1	2	$1 + 2 + 2 = 5$
1	3	$1 + 2 - 3 = 0$
2	1	$1 - 2 + 1 = 0$
2	2	$1 - 2 + 2 = 1$
2	3	$1 - 2 - 3 = -4$

Graphische Darstellung



Verlauf parallel
Abstand $\tau_1 - \tau_2$

Darstellung mit Referenz-Kodierung

$$Y = (e \quad Z_1(C) \ldots Z_{K_1-1}(C) Z_1(D) \ldots Z_{K_2-1}(D)) \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_{K_1-1} \\ \gamma_1 \\ \vdots \\ \gamma_{K_2-1} \end{pmatrix} + \epsilon \quad (4.7)$$

Interpretation:
τ_k, γ_l Abweichung von der Referenzkategorie

Beispiel: Designmatrix

Modell mit einem zweistufigen und einem dreistufigen Faktor und jeweils zwei Beobachtungen pro Faktorkombination

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$b)$ Modell mit Interaktion

Interaktionen lassen sich durch Aufnahme aller Produktterme $Z_k^e(C)Z_l^e(D)$ modellieren:

$$E(Y) = (e, Z_1^e(C)...Z_{K_2-1}(D), Z_1^e(C)Z_1^e(D)...Z_{K_1-1}(C)Z_{K_2-1}(D)) \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \gamma_{K_2-1} \\ (\tau\gamma)_{11} \\ \vdots \\ (\tau\gamma)_{K_1-1,K_2-1} \end{pmatrix}$$

Test auf Interaktion:

$$H_0 : (\tau\gamma)_{11} = \ldots = (\tau\gamma)_{K_1-1,K_2-1} = 0$$

Beispiel

Design-Matrix X für 2-Faktor Modell mit einem zweistufigen und einem dreistufigen Faktor: (jeweils eine Beobachtung pro Merkmalskombination).

$$X\beta = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \mu \\ \tau_1 \\ \gamma_1 \\ \gamma_2 \\ (\tau\gamma)_{11} \\ (\tau\gamma)_{12} \end{pmatrix}$$

Beispiel: 2 kategoriale Einflussgrößen mit Interaktion

$\mu = 1$

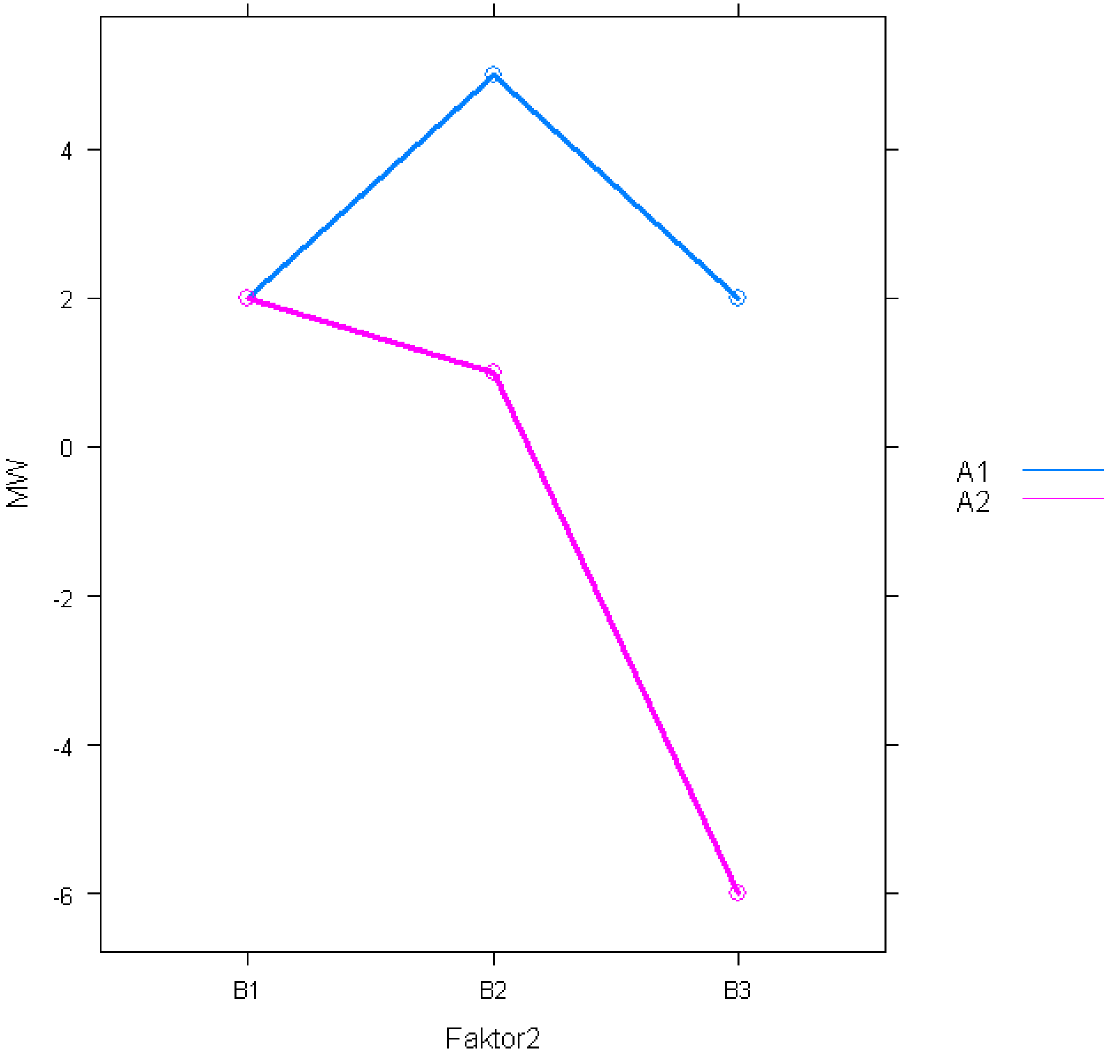
Faktor A (2-stufig):	$\tau_1 = 2$	$\Rightarrow \tau_2 = -2$
Faktor B (3-stufig):	$\gamma_1 = 1, \gamma_2 = 2$	$\Rightarrow \gamma_3 = -3$

Interaktion:	$(\tau\gamma)_{11} = -2$	$(\tau\gamma)_{21} = 2$
	$(\tau\gamma)_{12} = 0$	$(\tau\gamma)_{22} = 0$
	$(\tau\gamma)_{13} = 2$	$(\tau\gamma)_{23} = -2$

Berechnung der Mittelwerte:

Faktor A	Faktor B	MW
1	1	$1 + 2 + 1 - 2 = 2$
1	2	$1 + 2 + 2 + 0 = 5$
1	3	$1 + 2 - 3 + 2 = 2$
2	1	$1 - 2 + 1 + 2 = 2$
2	2	$1 - 2 + 2 + 0 = 1$
2	3	$1 - 2 - 3 - 2 = -6$

Graphische Darstellung



Verlauf unterschiedlich
$\Leftrightarrow$ Interaktion

Erweiterung auf Kombination von diskreten und stetigen Merkmalen (Kovarianzanalyse)

Beispiel für Design-Matrix X für $K = 3$ Gruppen mit je $n_k = 2$ Beobachtungen pro Gruppe und stetigem Merkmal x:

$$X = \begin{pmatrix} 1 & 0 & 0 & x_1 \\ 1 & 0 & 0 & x_2 \\ 0 & 1 & 0 & x_3 \\ 0 & 1 & 0 & x_4 \\ 0 & 0 & 1 & x_5 \\ 0 & 0 & 1 & x_6 \end{pmatrix} \beta = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_4 \end{pmatrix}$$

Interpretation:
In den drei Gruppen drei parallele Geraden mit Achsenabschnitt α_i und Steigung β_4

Erweiterung auf Geraden mit versch. Steigung

Modell:

$$Y_{kl} = \alpha_k + \beta_k X_{kl} + \varepsilon_{kl} \tag{4.8}$$

Matrixdarstellung (3 Gruppen, 2 Beobachtungen pro Gruppe)

$$X = \begin{pmatrix} 1 & 0 & 0 & x_1 & 0 & 0 \\ 1 & 0 & 0 & x_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & x_3 & 0 \\ 0 & 1 & 0 & 0 & x_4 & 0 \\ 0 & 0 & 1 & 0 & 0 & x_5 \\ 0 & 0 & 1 & 0 & 0 & x_6 \end{pmatrix} \beta = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Interaktion bedeutet Steigungen verschieden.

Test auf Interaktion: $\beta_1 = \beta_2 = \beta_3$

Darstellung mit Referenzkodierung

Modell:

$$
\begin{aligned}
Y_{kl} &= \alpha_3 + \alpha_k + \beta_3 X_{kl} + \beta_k X_{kl} + \varepsilon_{kl} \qquad (k = 1, 2) \\
Y_{kl} &= \alpha_3 + \beta_3 X_{kl} + \varepsilon_{kl} \qquad (k = 3)
\end{aligned}
$$

Matrixdarstellung (3 Gruppen 2 Beobachtungen pro Gruppe)

$$
X = \begin{pmatrix} 1 & 1 & 0 & x_1 & x_1 & 0 \\ 1 & 1 & 0 & x_2 & x_2 & 0 \\ 1 & 0 & 1 & x_3 & 0 & x_3 \\ 1 & 0 & 1 & x_4 & 0 & x_4 \\ 1 & 0 & 0 & x_5 & 0 & 0 \\ 1 & 0 & 0 & x_6 & 0 & 0 \end{pmatrix} \beta = \begin{pmatrix} \alpha_3 \\ \alpha_1 \\ \alpha_2 \\ \beta_3 \\ \beta_1 \\ \beta_2 \end{pmatrix}
$$

Interaktion bedeutet Steigungen verschieden.

Test auf Interaktion: $\beta_1 = \beta_2 = 0$

Typen von Quadratsummen

In der Literatur unterscheidet man 3 Typen von Quadratsummen:

Typ1 Sequentielle Quadratsummen

Typ2 Partielle Quadratsummen ohne höhere Interaktionsterme

Typ3 Partielle Quadratsummen

Beispiel mit Effekten X1 X2 X3 X1*X2 X1*X3 X2*X3 X1*X2*X3

	Variablen im Modell bei		
Effekt	Typ1	Typ3	Typ2
X1	-	alle anderen	X2 X3 X2*X3
X2	X1	alle anderen	X1 X3 X1*X3
X1*X2	X1 X2 X3	alle anderen	X1 X2 X3 X1*X3 X2*X3
X1*X2*X3	alle anderen	alle anderen	alle anderen

Beachte: Falls Interaktionen vorhanden sind, sind die einfachen Koeffizienten schwer direkt interpretierbar!

Beispiel: Vorher-Nachher Vergleich

X1 : Blutwert zum Zeitpunkt 1
X2 : Blutwert zum Zeitpunkt 2
Z : Gruppenzugehörigkeit (0= Placebo / 1= Verum)
oder auch allgemeine Gruppierungsvariable

Fragestellung: Gibt es einen Unterschied zwischen den Gruppen ?

Variante 1: Betrachte Differenzen X2-X1 = D und
führe 2 Stichproben t-Test durch
Vorteil: Einfach und leicht interpretierbar

Variante 2: Regression $X2 = \beta_0 + \beta_1 X1 + \gamma * Z + \epsilon$
Teste die Nullhypothese: $H_0 : \gamma = 0$
Vorteil: „Regression to the mean“ und mögliche Abhängigkeit der Differenz vom Anfangswert wird berücksichtigt.