SPM-Kurs 2011

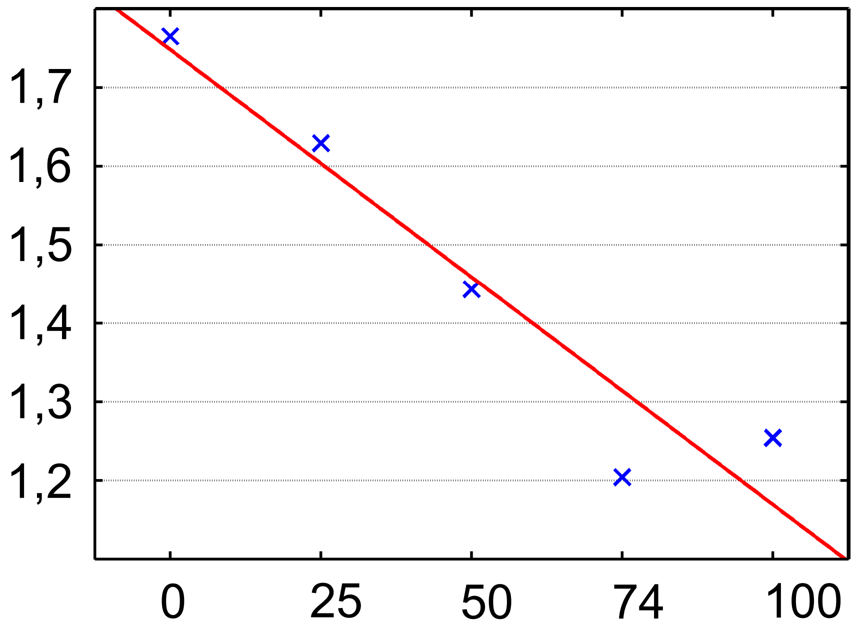
Allgemeines Lineares Modell
General Linear Model

Tobias Sommer-Blöchl,
NeuroImage Nord,
Institut für Systemische Neurowissenschaften

Universitätsklinikum
Hamburg-Eppendorf

- GLM: (multiple) Regression $y = ax + b + Fehler$
- T-Tests und Varianzanalysen: Spezialfall des GLM
- abhängige Kriteriumsvariable
- unabhängige Prädiktorvariable/Regressor

x	y
0	1,77
25	1,63
50	1,44
75	1,20
100	1,25



Ziel: ⇒ Beschreibung des Zusammenhangs zwischen x und y

- bei mehreren Regressoren x_j: $y = a_1 x_1 + a_2 x_2 + \ldots + a_m x_m + 1b + \varepsilon$

⇒ Kriterium y als gewichtete Linearkombination der Regressoren x_j

GLM in Matrizenschreibweise

$$y = a_1 x_1 + a_2 x_2 + \ldots + a_m x_m + 1b + \varepsilon$$

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}}_{\text{Kriteriumsvariable}} = \overbrace{\beta_1 \begin{bmatrix} x_{1,1} \\ \vdots \\ x_{i,1} \\ \vdots \\ x_{n,1} \end{bmatrix} + \beta_2 \begin{bmatrix} x_{1,2} \\ \vdots \\ x_{i,2} \\ \vdots \\ x_{n,2} \end{bmatrix} + \ldots + \beta_m \begin{bmatrix} x_{1m} \\ \vdots \\ x_{im} \\ \vdots \\ x_{nm} \end{bmatrix}}^{\text{gewichtete Regressoren}} + \overbrace{\beta_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}}^{\text{Konstante}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\text{Fehler}}$$

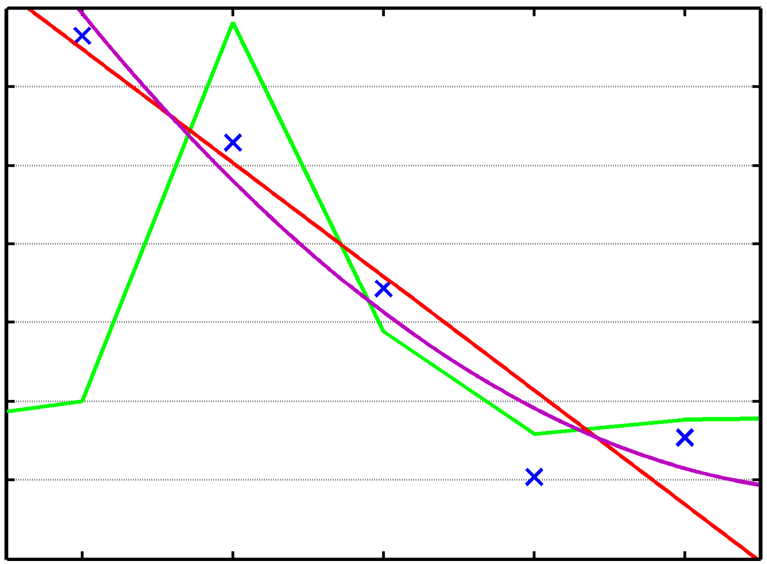
$$\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} & x_{10} \\ \vdots & \vdots & & \vdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{im} & x_{i0} \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} & x_{n0} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$y = X\beta + \varepsilon$$

Regressionsanalyse

1. **Modellauswahl**
2. **Parameterschätzung**
3. **Beurteilung der Modellgüte**
4. **Inferenz/Signifikanztestung**

Modellauswahl

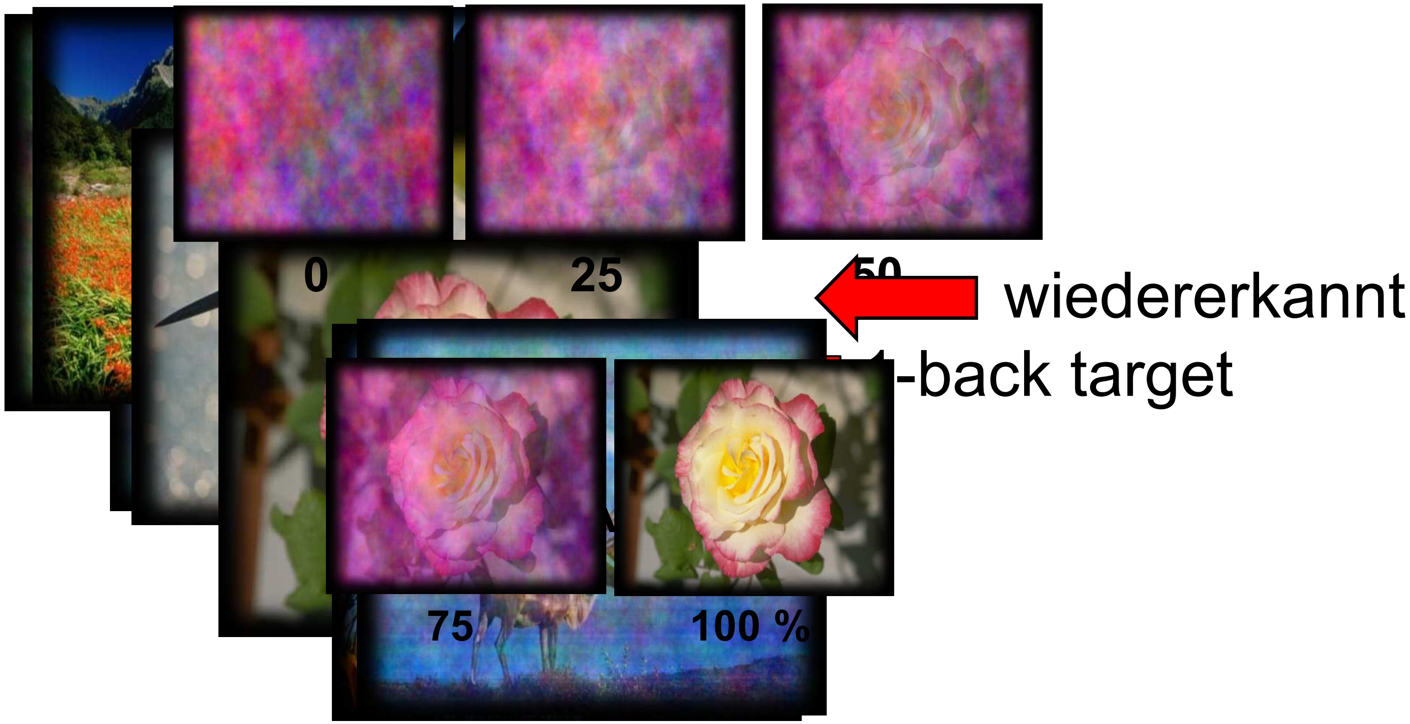


- Welche Art von Zusammenhang wird hypothetisiert?
- Welche Regressoren werden in das Modell aufgenommen?

Beispielstudie:

- Enkodierung: Bilder unterschiedlich gut sichtbar

  Aufgabe: n-back



- Gedächtnistest: alle Bilder voll sichtbar

  Aufgabe: alt/neu-Entscheidung

Beispielstudie

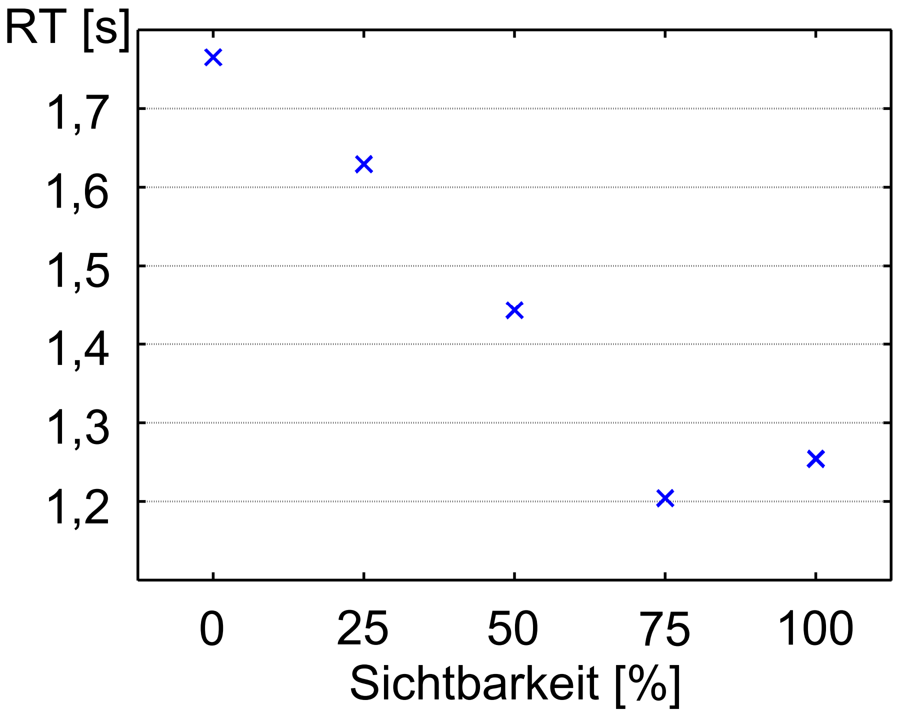
⇒ Kriteriumsvariable:

mittlere Reaktionszeit in Gedächtnistest

⇒ 1 Regressor: + Konstante (y-Achsenabschnit)

Sichtbarkeit in % der Bilder während des Enkodierens

⇒ Annahme eines linearen Zusammenhangs zwischen Sichtbarkeit und Reaktionszeit



$$y = \beta_1 x + 1\beta_0 + \varepsilon$$

$$y = X\beta + \varepsilon$$

$$y = X\beta + \varepsilon$$

y	x
1,77	0
1,63	25
1,44	50
1,20	75
1,25	100

x_1 : Sichtbarkeit

$$\text{Reaktionszeiten} \begin{bmatrix} 1{,}77 \\ 1{,}63 \\ 1{,}44 \\ 1{,}20 \\ 1{,}25 \end{bmatrix} = \begin{bmatrix} \overbrace{0}^{x_1} & 1 \\ 25 & 1 \\ 50 & 1 \\ 75 & 1 \\ 100 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

β_0 : Y-Achsenabschnitt

β_1 : Steigung

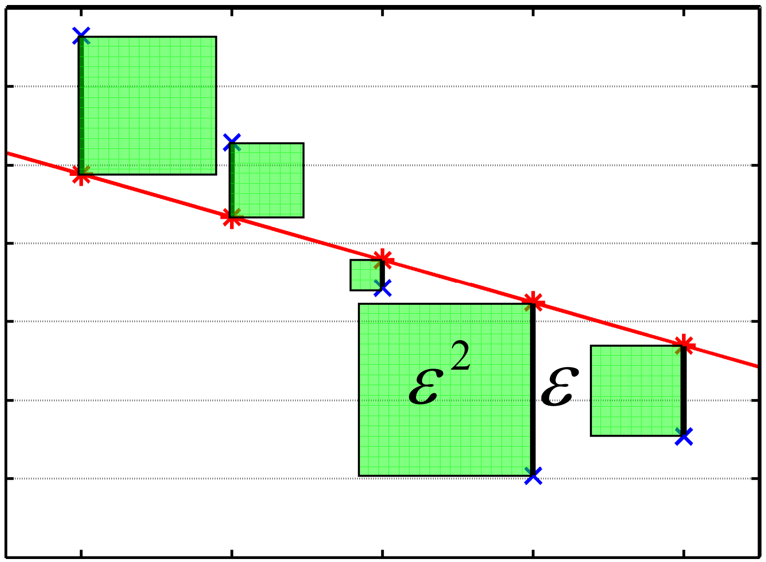
Parameterschätzung

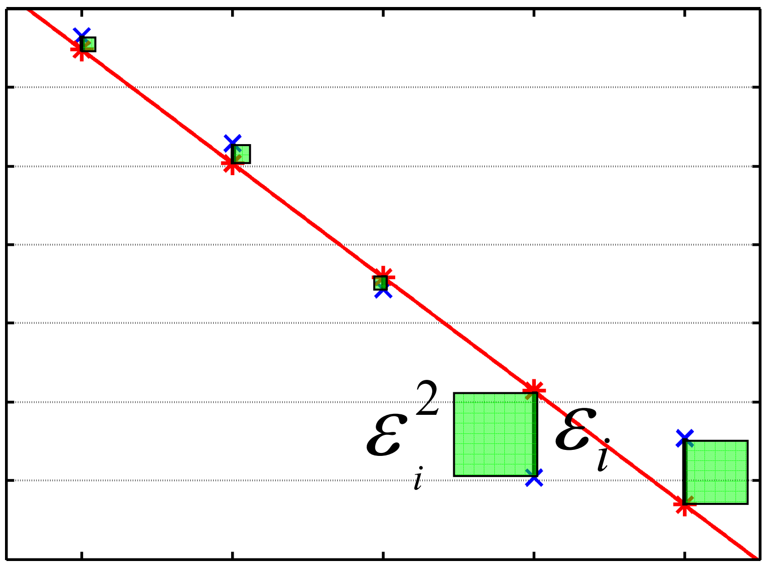
Ziel:

- Regressionsgewichte bestimmen, so dass die Kriteriums-variable möglichst genau vorhergesagt wird
- Fehler soll minimiert werden

⇒ **Kriterium der kleinsten Abweichungsquadrate:**

- **SAQ** (Summe der Abweichungsquadrate) soll minimiert werden





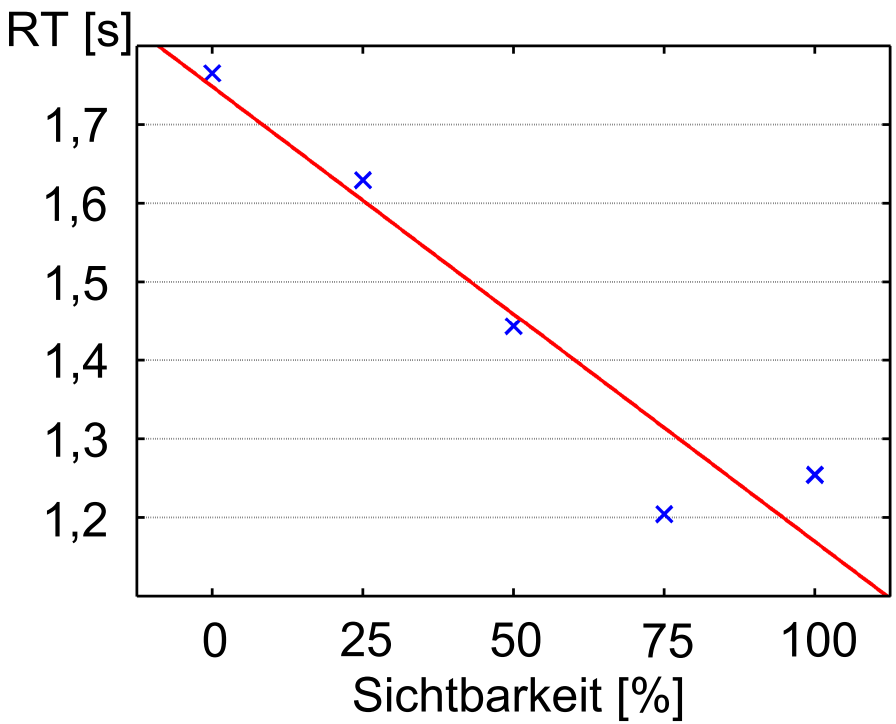
Beispielstudie:

• 1 Regressor + 1 Konstante

⇒ 2 Regressionsgewichte/Parameter zu schätzen

$$\beta = (X'X)^{-1} X'y = \begin{bmatrix} -0{,}006 \\ 1{,}752 \end{bmatrix}$$

$$y = -0{,}006\,x + 1{,}75 + \varepsilon$$



$$SAQ = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon$$

$$y = X\beta + \varepsilon$$
$$\varepsilon = y - X\beta$$

$$= (y - X\beta)'(y - X\beta)$$

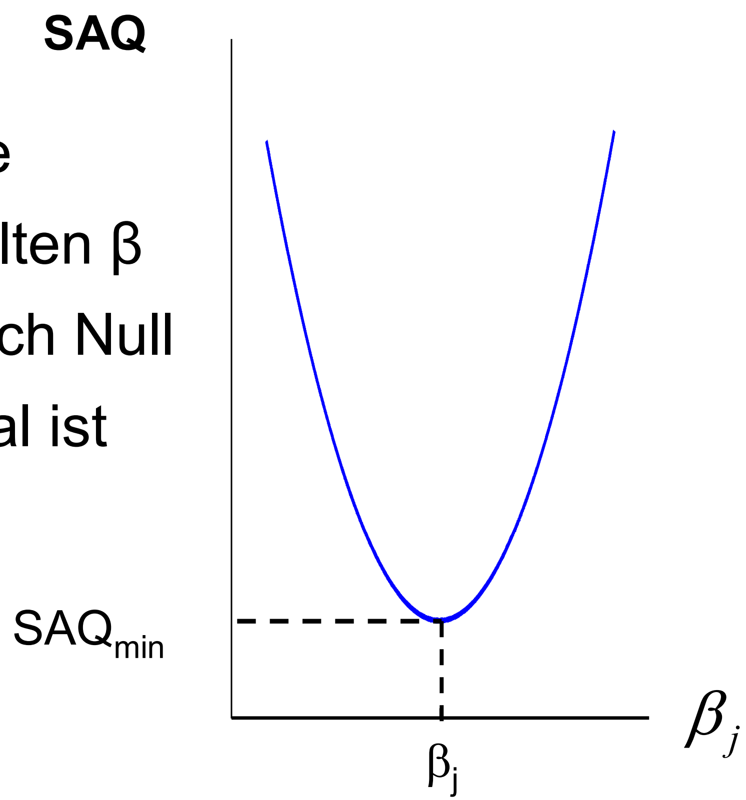
$$= y'y - 2\beta'X'y + \boxed{\beta'X'X\beta}$$

$\Rightarrow$ quadratische Gleichung ($y=x^2$)

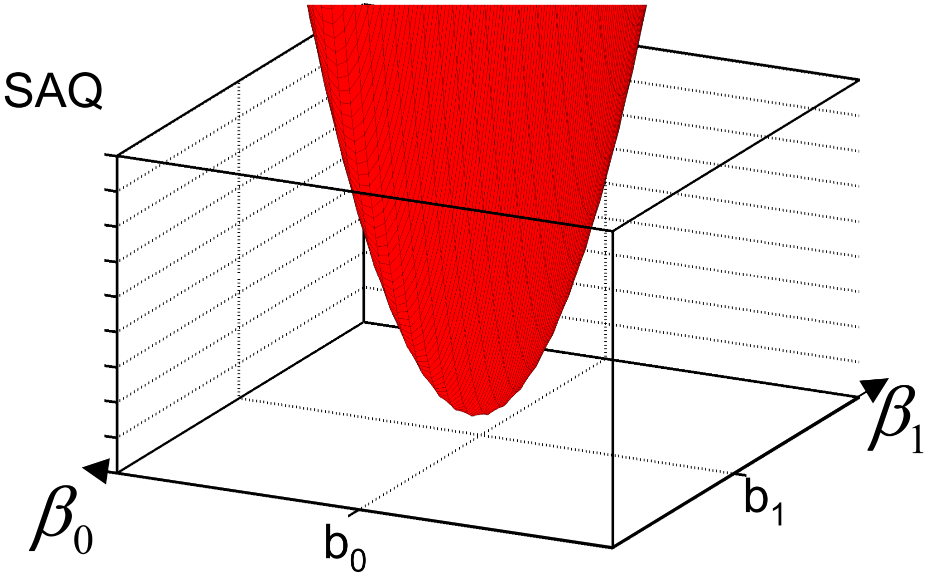
- bei einem Parameter bildet die SAQ eine Parabel in Abhängigkeit von dem gewählten β
- durch die partielle Ableitung der SAQ nach Null findet man das β, für die die SAQ minimal ist

$$\frac{\partial SAQ}{\partial \beta'} = -2X'y + 2X'X\beta = 0$$

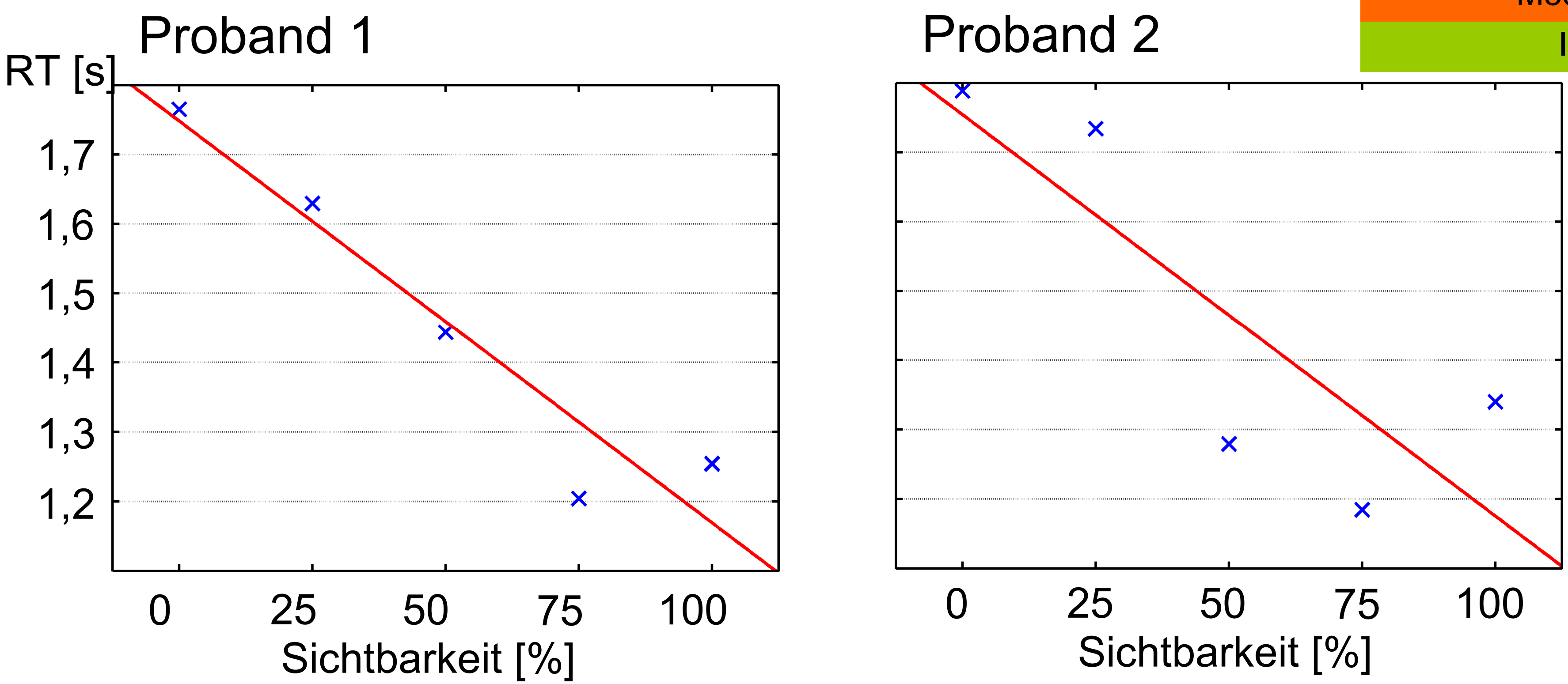
$$\beta = (X'X)^{-1} X'y$$



- bei zwei Parametern, wie im Beispiel (1 Regressor und Konstante), bilden die SAQ somit in Abhängigkeit von der Größe der beiden β s einen dreidimensionalen Kegel mit genau einem Minimum
- durch die partielle Ableitung nach Null erhält man die beiden β s, an denen die SAQ minimal ist
- Parameterschätzung für alle Parameter in diesem *Ordinary Least Square* (OLS) Verfahren in einem Schritt



Beurteilung der Modellgüte



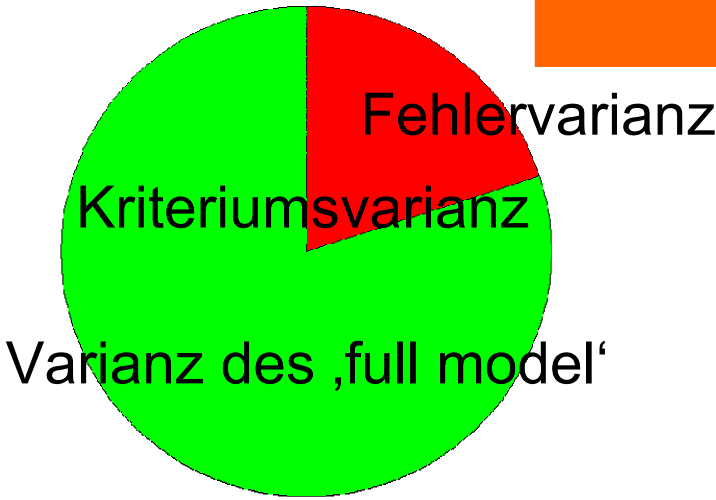
$$y = -0{,}006\,x + 1{,}75 + \varepsilon$$

Welchen Anteil der beobachteten Varianz erklärt das Modell?

Begriffsbestimmung:

i) uneingeschränktes oder ‚**full**' Modell:

$$y = \beta_1 x_1 + \beta_0 1 + \varepsilon$$



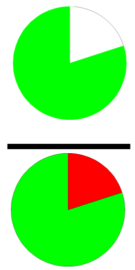
ii) eingeschränktes oder ‚**reduced**' Modell

$$y = \beta_0 1 + \varepsilon$$

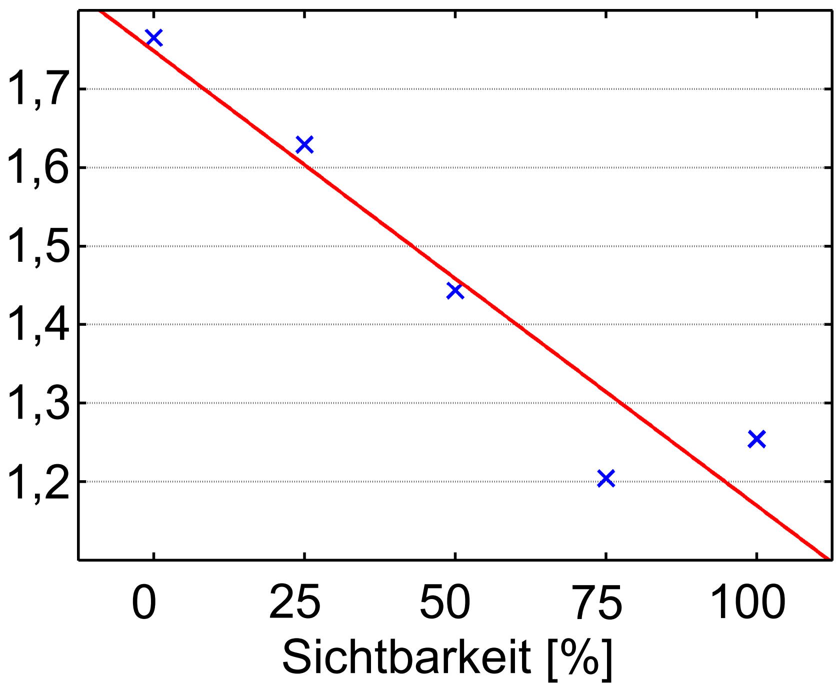
Determinationskoeffizient

Welchen Anteil der beobachteten Varianz erklärt das Modell?

$$R^2 = \frac{\textit{erklärte Varianz}}{\textit{beobachtete Varianz}} =$$

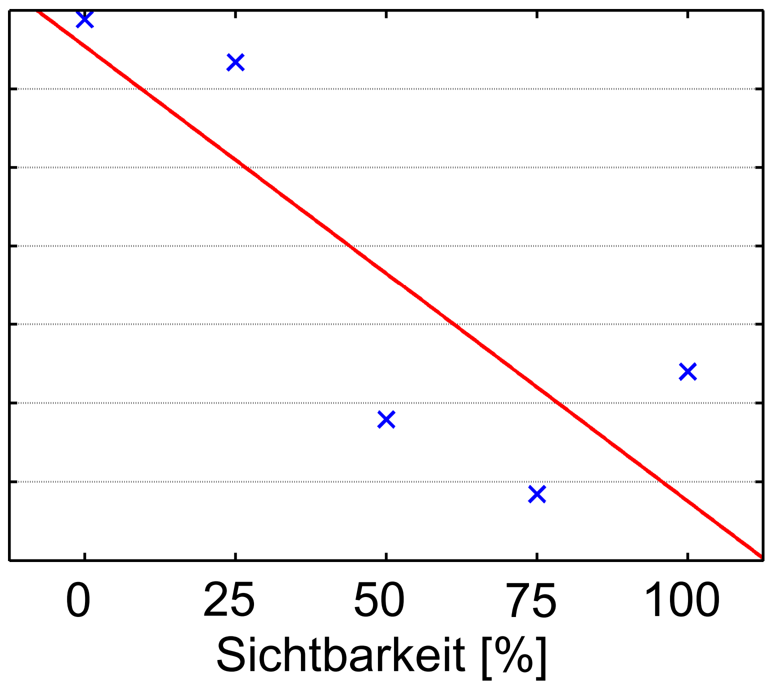


Proband 1



$$R^2 = 0{,}91$$

Proband 2



$$R^2 = 0{,}39$$

- SPM: Modellgüte durch Residuen/Fehler (ResMs.img) oder über F-Test (‘effects of interest’) bestimmbar

Inferenz

Ausgangspunkt

- Model ausgewählt
- Regressionsgewichte/Parameter bestimmt
- Determinierte Varianz bestimmt

Ziel der Inferenz

- Überprüfung einer Hypothese anhand von Stichprobendaten

I) Formulierung der Hypothesen

i) Nullhypothese H_0: Hypothese, die man verwerfen möchte

ii) Alternativhypothese H_1: Hypothese, die man belegen möchte

- t-Test: Mittelwerts- vs. F-Test: Varianzen-Unterschiede
- „t-Test-Hypothesen" in fMRT häufiger

II) Festlegung der Signifikanzschwelle/Irrtumswahrscheinlichkeit: α

- Wahrscheinlichkeit, mit der ein Ergebnis **nicht** zufällig zustande kam
- Wahrscheinlichkeit, einen Fehler 1. Art zu begehen

Regressionsanalyse

Modellauswahl

Parameterschätzung

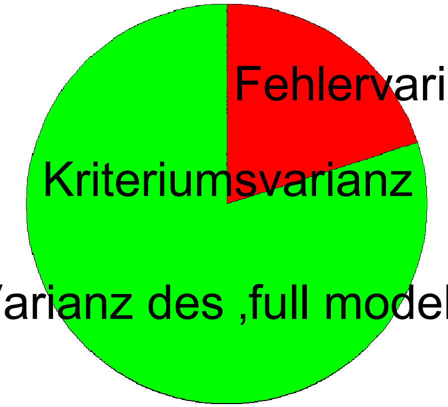
Modellgüte

Inferenz

III) Berechnung des F-Werts

- F-Wert ist das Verhältnis zweier Varianzen

‚full model‘ ↔ **‚reduced model‘**



$$F = \frac{\text{Varianz ‚full model‘}}{\text{Fehlervarianz ‚full model‘}}$$

$$F = \frac{(\text{Fehlervarianz ‚red. model‘}) - (\text{Fehlervarianz ‚full model‘})}{\text{Fehlervarianz ‚full model‘}}$$

IV) Freiheitsgrade

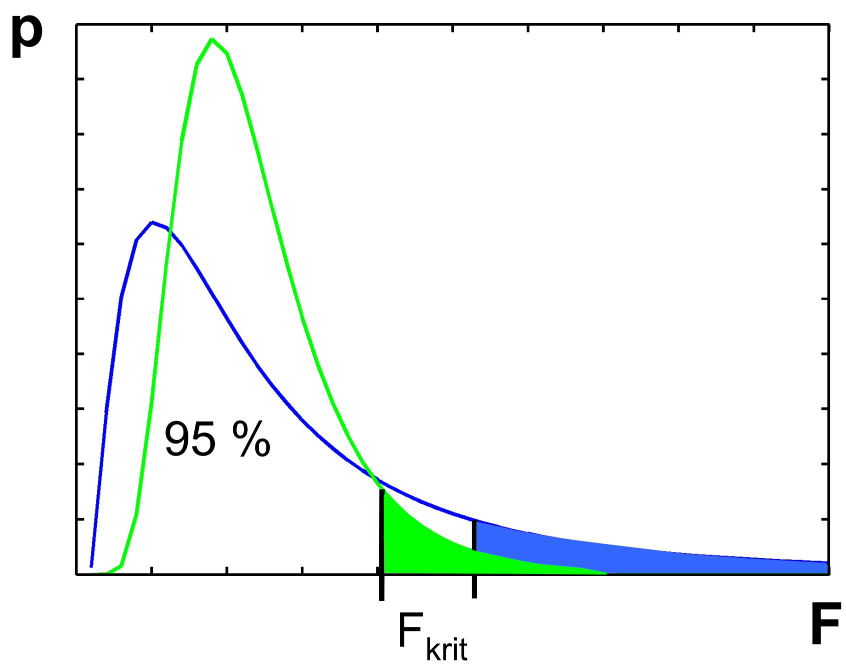
- F-Wert ist das Verhältnis zweier Varianzen
- dividiert durch die Freiheitsgrade (df)

$$F = \frac{\left[\text{Varianz 1} \,/\, df_{\text{Zähler}}\right]}{\left[\text{Varianz 2} \,/\, df_{\text{Nenner}}\right]}$$

- Berechnung der Freiheitsgrade:
  - $df_{\text{Fehlervarianz}}$ = Anzahl Beobachtungen – Anzahl Regressoren
  - $df_{\text{Varianz full model}}$ = Anzahl Regressoren
- Einfluss auf F-Wert:
  - Anzahl der Regressoren $\uparrow$ $\Rightarrow$ F-Wert $\downarrow$
  - Anzahl der Beobachtungen $\uparrow$ $\Rightarrow$ F-Wert $\uparrow$

V) F-Verteilung

- Wahrscheinlichkeitsverteilung



F(5,5) = blau

F(20,20) = grün

- $F > F_{krit}$: Wahrscheinlichkeit, dass das ‚**full model**' nur zufällig mehr Varianz erklärt als das ‚**reduced model**' ist kleiner als $\boldsymbol{\alpha}$
- H_0 wird verworfen, H_1 angenommen

Beispielstudie:

⇒ **Formulierung der Hypothesen**

i) Nullhypothese H_0:

Es gibt **keinen** Zusammenhang zwischen der Sichtbarkeit der Bilder während des Enkodierens und den Reaktionszeiten im Gedächtnistest.

ii) Alternativhypothese H_1:

Es gibt **einen signifikanten** Zusammenhang.

⇒ **Festlegung der Signifikanzschwelle auf $\alpha = 0{,}05$**

Die H_0 wird verworfen, wenn die Wahrscheinlichkeit, dass das Ergebnis zufällig zustande kam kleiner als 5 % ist.

⇒ Berechnung des F-Wertes

$$F = \frac{\left[\text{Varianz 1} \,/\, df_{\text{Zähler}} \right]}{\left[\text{Varianz 2} \,/\, df_{\text{Nenner}} \right]}$$

- 1 $df_{\text{Zähler}}$: 1 Regressor und Konstante
- 4 df_{Nenner}: 5 Beobachtungen – 1 Regressor

$$F = \frac{\left[0{,}04 \,/\, 1 \right]}{\left[0{,}004 \,/\, 4 \right]} = F(1{,}4) = 40{,}4$$

- $F_{krit} = 21{,}2$

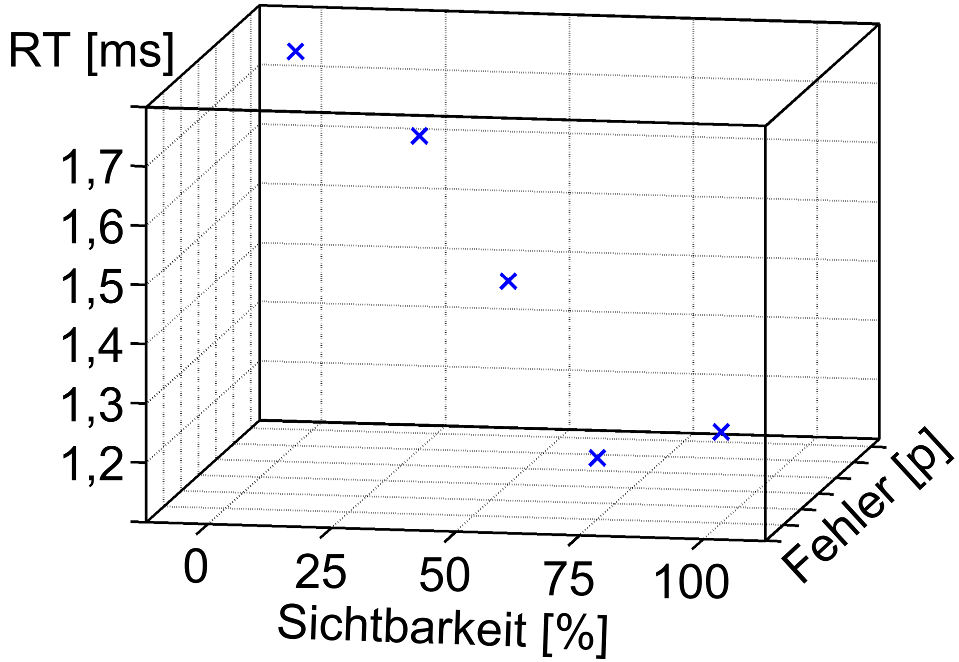
⇒ H_0 wird verworfen, H_1 wird angenommen

⇒ signifikanter Zusammenhang zwischen Sichtbarkeit der Bilder und der Reaktionszeit im Gedächtnistest bei Proband 1

Verbesserung des Modells

- ein weiterer Regressor:
    - Fehlerrate in der jeweiligen Sichtbarkeitsstufe

y	x_1	x_2
1,77	0	0,5
1,63	25	0,5
1,44	50	0,3
1,20	75	0,1
1,25	100	0,1

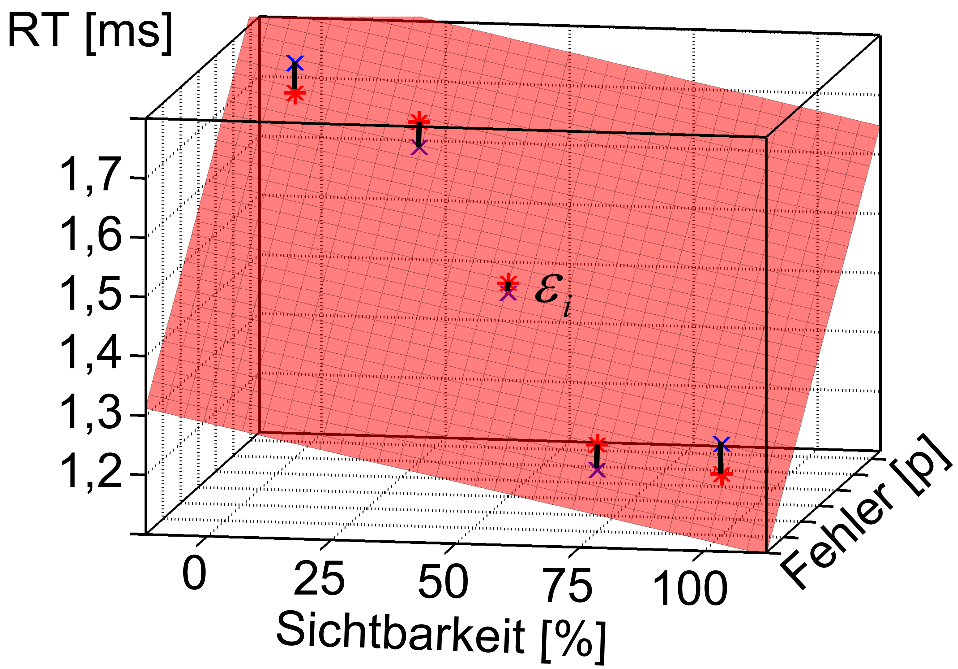


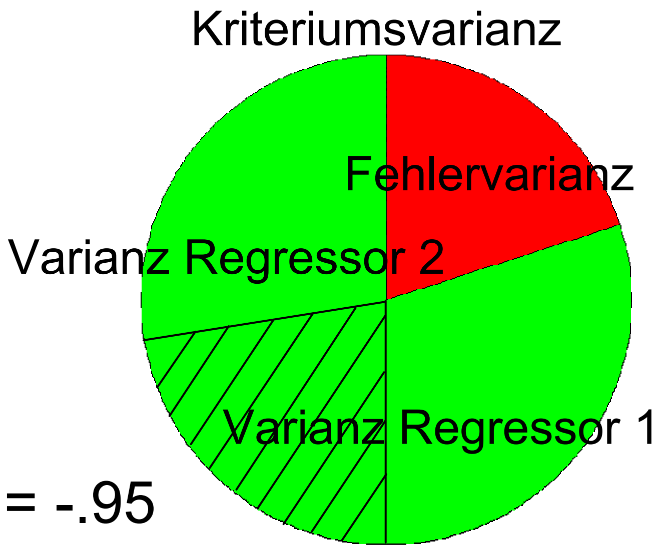
$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_0 1 + \varepsilon$$

$$\underset{\text{Reaktionszeiten}}{\begin{bmatrix} 1{,}765 \\ 1{,}629 \\ 1{,}443 \\ 1{,}204 \\ 1{,}254 \end{bmatrix}} = \begin{array}{c} \begin{array}{ccc} \text{Sichtbarkeit} & \text{Fehlerrate} & \\ x_1 & x_2 & \end{array} \\ \begin{bmatrix} 0 & 0{,}5 & 1 \\ 25 & 0{,}5 & 1 \\ 50 & 0{,}3 & 1 \\ 75 & 0{,}1 & 1 \\ 100 & 0{,}1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\boldsymbol{b} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y} = \begin{bmatrix} -0{,}002 \\ 0{,}85 \\ 1{,}3 \end{bmatrix}$$

$$y = -0{,}002x_1 + 0{,}85x_2 + 1{,}3 + \varepsilon$$





- Determinationskoeffizient: $R^2 = 0{,}96$
- **Inkrement**: 5% zusätzlich erklärte Varianz
- die beiden Regressoren sind hoch korreliert r = -.95

  $\Rightarrow$ Suppressionseffekt:

  - β_1 im ersten Modell 1,4, im zweiten nur 0,4

$\Rightarrow$ F(2,3) = 36,9 > F_{krit} = 30,8

Varianzanalyse

• Varianzanalyse: Spezialfall des GLM

• T-Tests: Spezialfall der Varianzanalyse

⇒ gleiche Formeln wie bei Regressionsanalyse

⇒ aber: i) Kodierung der Daten

ii) Design-Matrix

Kodierung

i) Voraussetzung für GLM sind intervallskalierte Daten

ii) Einzelvergleiche einzelner Faktorstufen

⇒ verschiedene Arten der Kodierung

⇒ SPM: Dummykodierung mit Konstante

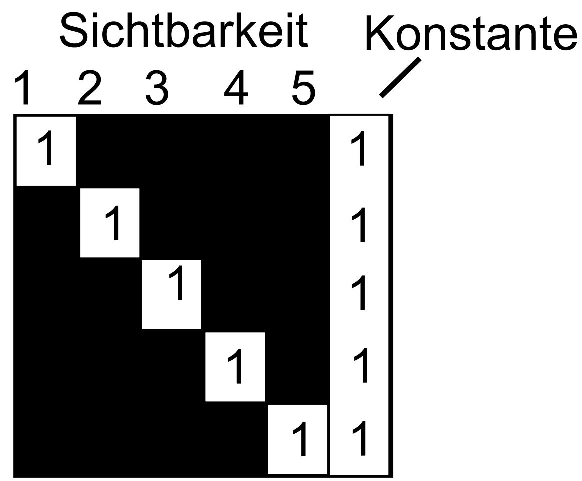
Beispielstudie:

- gleiches Model und Daten wie in der Regressionsanalyse:
  - Kriteriumsvariable: Reaktionszeit im Gedächtnistest
  - Regressoren: Sichtbarkeit der Bilder kodiert als 5 Faktorstufen

y	x
1,77	0
1,63	25
1,44	50
1,20	75
1,25	100

Kodierung $\Rightarrow$

$$
\begin{array}{cccccc}
\multicolumn{5}{c}{\text{Sichtbarkeit}} & \text{Konstante} \\
1 & 2 & 3 & 4 & 5 & \\
\end{array}
\\
\underbrace{\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}}_{\text{Designmatrix X}}
=
$$



Visualisierung der Designmatrix X

$\Rightarrow$ **Grenzen des Beispielmodels**:

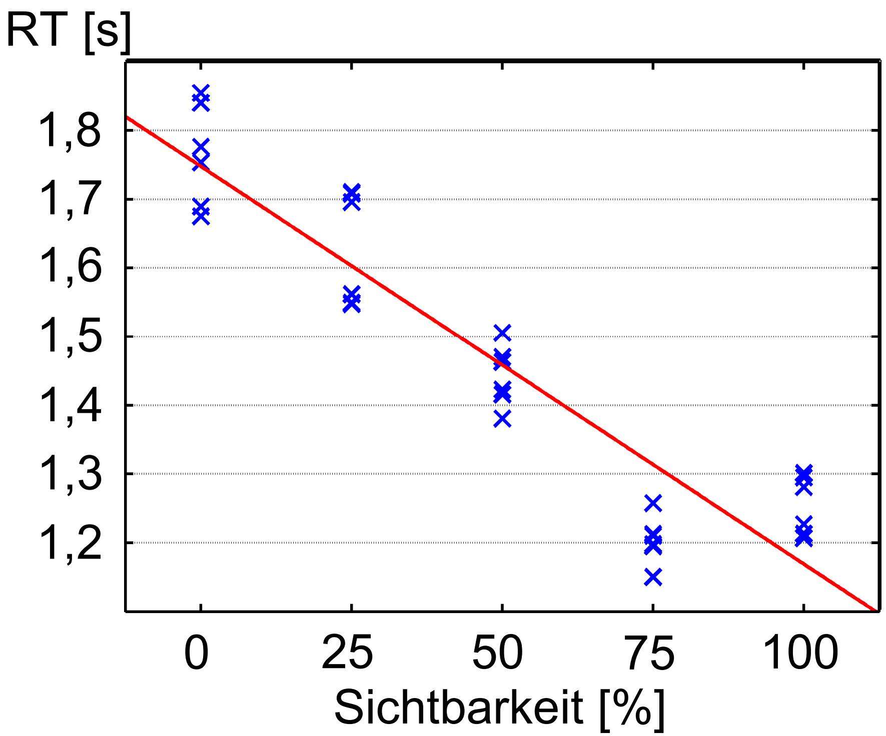
- keine Varianz in den einzelnen Faktorstufen
- Varianzanalyse nicht möglich

Modell:

- bisher: Mittelwerte der Reaktionszeit einer Person
- jetzt: einzelne Trials (n=30) im Gedächtnistest

$\Rightarrow$ **Regressionsanalyse**

y	x
1,43	50
1,58	25
1,45	25
1,85	0
1,06	100
1,59	0
1,37	75
1,46	50
1,33	75
1,21	100
..	...



$$\begin{bmatrix} 1{,}43 \\ 1{,}58 \\ 1{,}45 \\ 1{,}85 \\ 1{,}06 \\ 1{,}59 \\ 1{,}37 \\ 1{,}46 \\ 1{,}33 \\ 1.21 \\ \vdots \end{bmatrix} = \begin{bmatrix} 50 & 1 \\ 25 & 1 \\ 25 & 1 \\ 0 & 1 \\ 100 & 1 \\ 0 & 1 \\ 75 & 1 \\ 50 & 1 \\ 75 & 1 \\ 100 & 1 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$b = (X'X)^{-1} X'y = \begin{bmatrix} -0{,}006 \\ 1{,}75 \end{bmatrix}$$

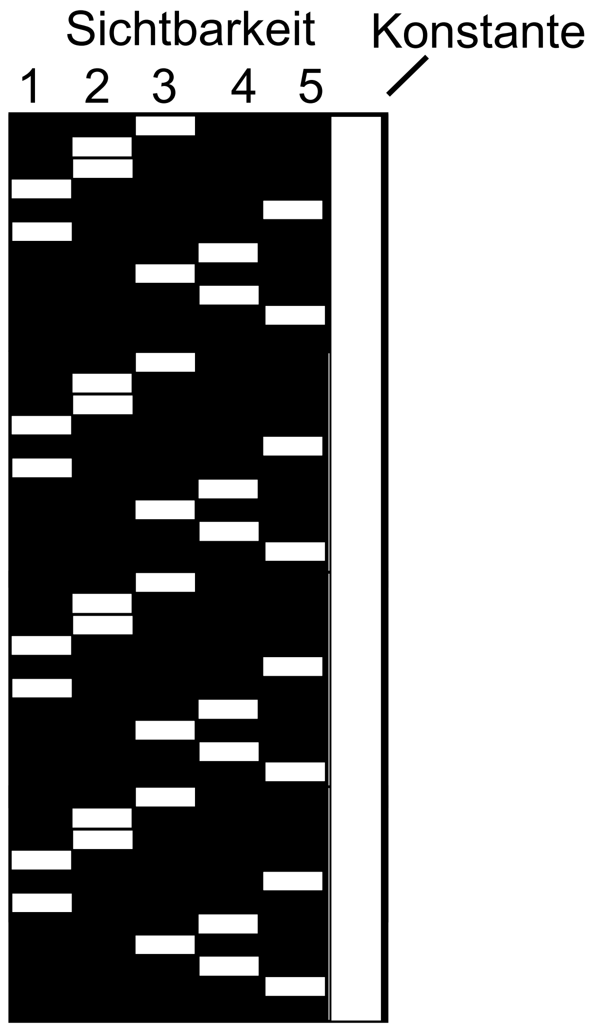
⇒ Varianzanalyse

y	x
1,43	50
1,58	25
1,45	25
1,85	0
1,06	100
1,59	0
1,37	75
1,46	50
1,33	75
1,21	100
..	...

Kodierung

Sichtbarkeit Konstante

$$
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & \\
\end{array}
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{bmatrix}
=
$$



⇒ Modellauswahl

- Erstellung der Designmatrix

⇒ Parameterschätzung

$$b = \left(X'X\right)^{-1} X'y = \begin{bmatrix} 0{,}55 \\ 0{,}41 \\ 0{,}23 \\ -0{,}01 \\ 0{,}03 \\ 1{,}25 \end{bmatrix}$$

⇒ Modellgüte

- $R^2 = 0{,}99$

(identisch zum regressionsanalytischen Modell)

Unterschiede von Design-Matrizen

„full-rank“

$$\begin{bmatrix} 0 & 0{,}5 & 1 \\ 25 & 0{,}5 & 1 \\ 50 & 0{,}3 & 1 \\ 75 & 0{,}1 & 1 \\ 100 & 0{,}1 & 1 \end{bmatrix}$$

„rank-deficient“

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$\Rightarrow$ Regressor als Linearkombination der anderen Regressoren

$$1 * x_1 + 1 * x_2 + 1 * x_3 + 1 * x_4 + 1 * x_5 = x_6$$

$\Rightarrow$ nicht alle Parameter können nicht eindeutig geschätzt werden

$\Rightarrow$ nicht alle Kontraste können geschätzt werden

(Kontrastgewichte müssen in der Summe 0 ergeben)

Unterschiede von Design-Matrizen

„full-rank“

$$\begin{bmatrix} 0 & 0{,}5 & 1 \\ 25 & 0{,}5 & 1 \\ 50 & 0{,}3 & 1 \\ 75 & 0{,}1 & 1 \\ 100 & 0{,}1 & 1 \end{bmatrix}$$

„rank-deficient“

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$b = \left(X'X\right)^{-1} X'y$$

Inverse von X

$$\beta = X^{+} y$$

Pseudoinverse von X
(in SPM)

GLM in der fMRT

-mass univariate approach:

- Aktivität in jedem Voxel wird unabhängig von jedem anderen Voxel analysiert
- (Gegensatz: multivariate approaches, e.g. pattern classification)
- erst am Ende wird die Aktivität in allen anderen Voxeln mitberücksichtigt und für multiple Vergleiche korrigiert

-abhängige Variable 'y' and Regressoren 'x':

- unterschiedlich in 1st and 2nd Level Analysen

1st Level oder Single Subject Analysen

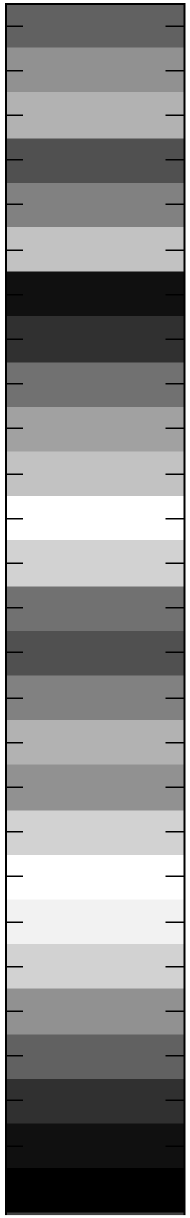
- abhängige Variable y:
  - Veränderung der BOLD-Signal Intensität in einem Voxel über die Zeitserie, also in einem Voxel in jedem Volumen über den Verlauf des Experiment

⇒ y = BOLD-Signal Intensität in einem Voxel über die Zeit

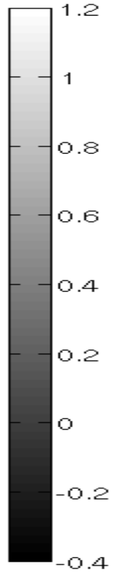
Zeitserie

Volumen	Signalintensität
1	0.2
2	0.5
3	0.7
4	0.1
5	0.4
6	0.8
7	-0.3
8	-0.1
9	0.3
10	0.6
11	0.8
12	1.2
13	0.9
14	0.3
15	0.1
16	0.4
17	0.7
18	0.5
19	0.9
20	1.2
21	1.1
22	0.9
23	0.5
24	0.2
25	-0.1
26	-0.3
27	-0.4

Grauwertskalierung



zufällige Skalierung

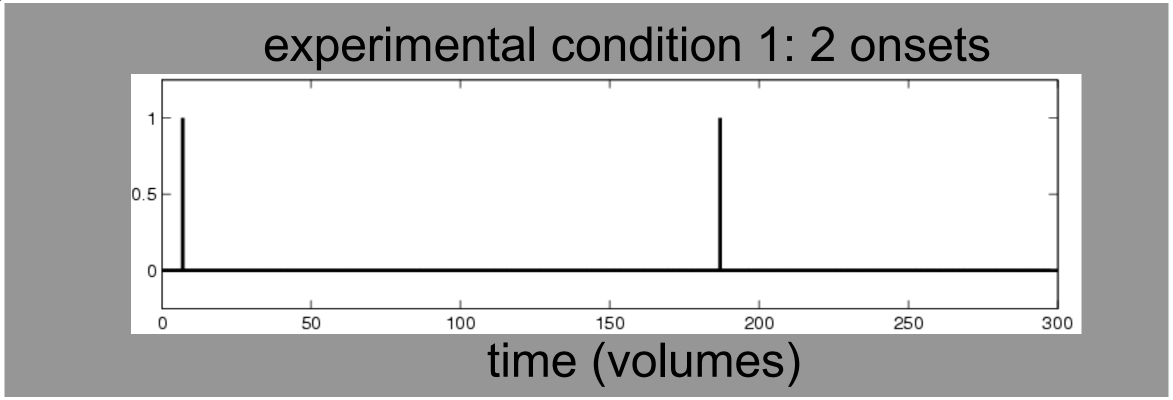


Maximum: weiß

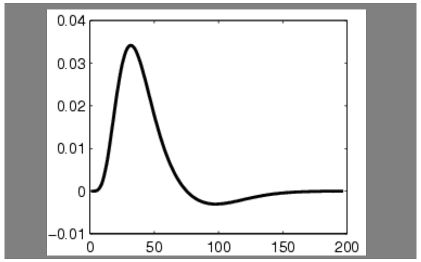
Minimum: schwarz

-Regressoren = unabhängige Variable:

- unabhänig – hängen vom Timing der experimentellen Bedingungen ab
- Neuronenverbände, die in die Prozessierung involviert sind, zeigen erhöhte synaptische Aktivität = sehr kurze Ereignisse
- δ-function (stick-function; 0 and 1) zu den Onsets der experimentellen Bedingungen



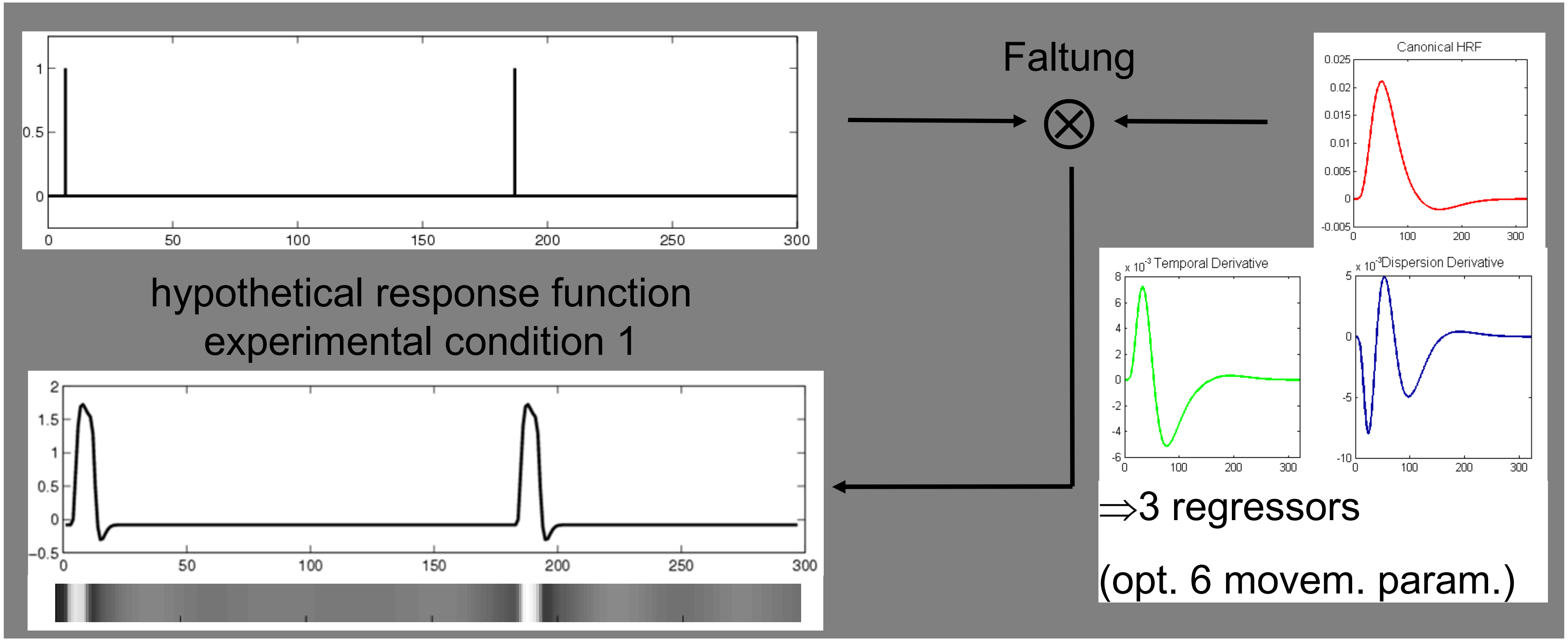
- neuronale Aktivität kann nicht direkt gemessen werden mittels fMRT
- neuronale Aktivität triggert eine **h**modynamic **r**esponse **f**unction (hrf)



-regressors:

– stick-function zu den Onsets der experimentellen Bedingungen

- triggert eine **h**modynamic **r**esponse **f**unction (hrf)



- Voxel mit Neuronenverbände, die in die Prozessierung involviert sind, sollten diese hypothetische hämodynamische Antwortfunktion zeigen ≈ BOLD Signal Veränderungen

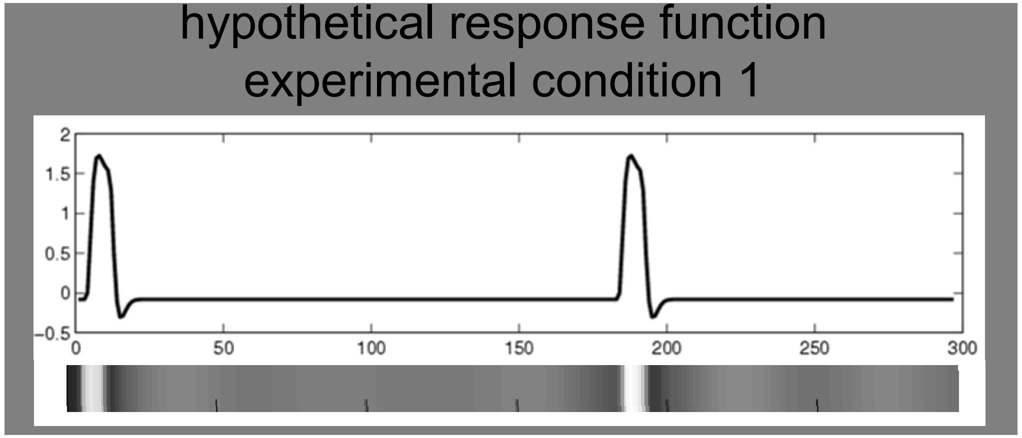
GLM in der fMRT – 1st Level

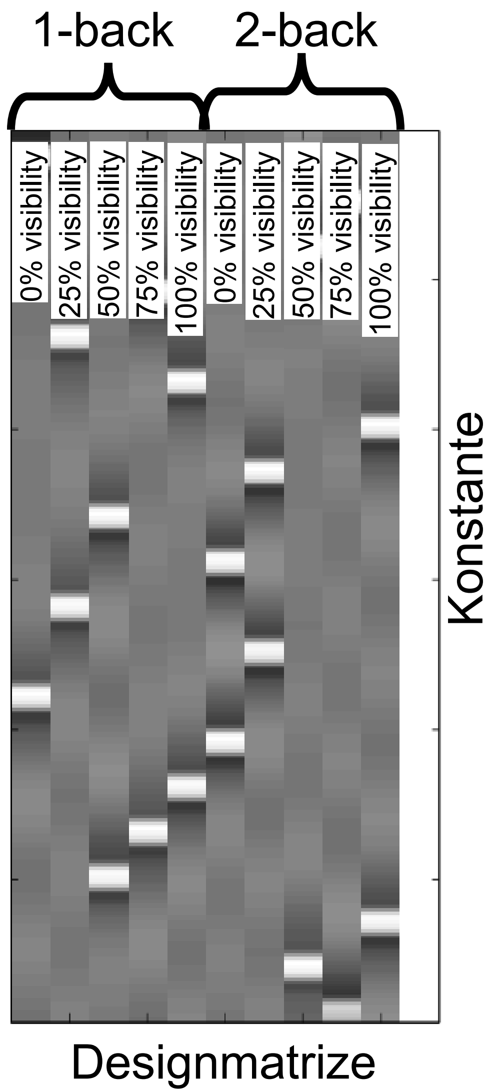
Modellauswahl

Parameterschätzung

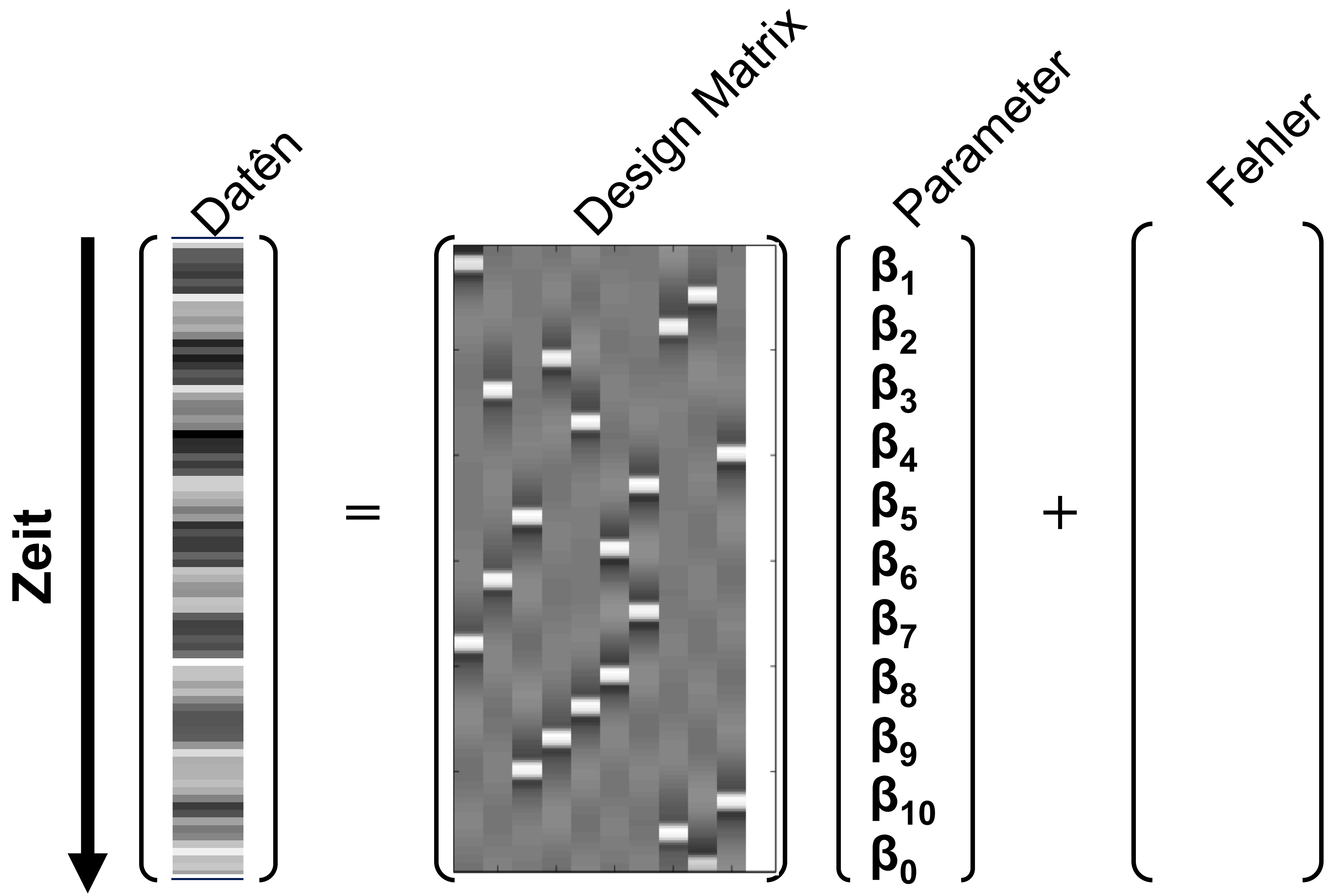
Modellgüte

Inferenz





GLM in der fMRT – 1^{st} Level



$$Y = X \bullet \beta + \varepsilon$$

$\Rightarrow$ Statistik immer für jedes Voxel unabhängig

2nd Level oder Gruppen-Analysen

- abhängige Variable y:
  - Ergebnis der 1st-Level Analysen in einem Voxel über die verschiedenen Versuchspersonen (Werte in den con-images)
- Regressor(en):
  - in 0 und 1 dummy-codierte Zuordnung dieser Ergebnisse zu den Bedingungen oder Faktoren, die auf dem 2nd-Level statistisch verglichen werden sollen