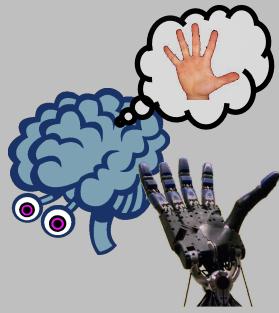


Group Assignments



| Mon 2/20 – Tue 2/21 9:00 AM 6:00 PM | Mon 2/27 – Tue 2/28 9:00 AM 6:00 PM | Mon 3/6 – Tue 3/7 9:00 AM 6:00 PM | Thu 3/9 – Fri 3/10 9:00 AM 6:00 PM |
|--|---|--|---|
| Frederick Thiemer Thomas Fessel Willy Robin Kinne Nastasia Pawlak Lillith Bitter Sakura Ihalage Alexander Vorwerg Mamaka Narava Anila Hung Nguyen | Astrid Günther Philo Reipke Andrea Weber Jasmin Seyfi Markus Flachmann Lorena Aldana Philipp Braukmann Beichert Thorben Drees Markus Vieth | Kyra Pritzel-Hentley Marie Klönne Jan-Hendrik Reese Manuel Klocke Justin Baudisch Matthias Barth Marvin Koppka Henrike Zimmermann Michel Bovender Kai Konen | Domenic Nagel Emanuel Dyck Marcel Friedrichs Haiko Ast Luisa Rajcsanyi Jens Loers Moritz Niebeling Katharina Kemmet Marcel Butzkies Moritz |

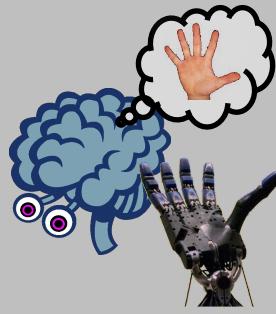
New Doodle Poll for Exams:

<http://doodle.com/poll/pxwx9mybi43mp36t> → “First come, first serve”

Fr 03.03.
Mo 13.03.
Tue 14.03

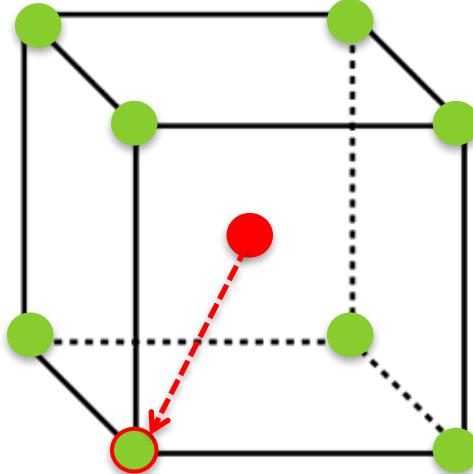
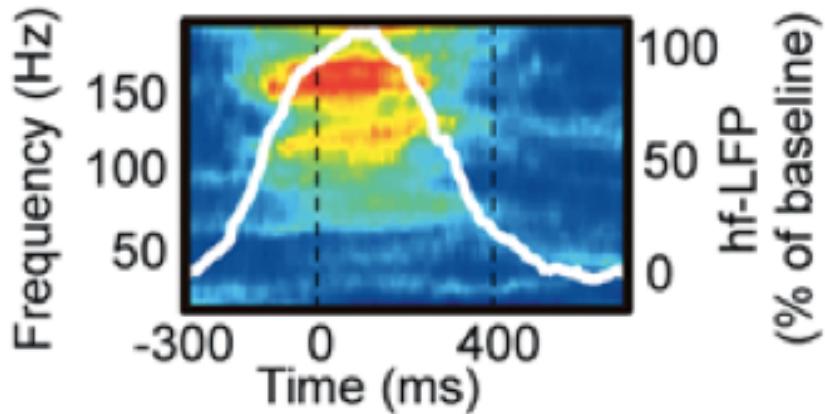
Fr 17.03
Mo 20.03.
Tue 21.03.

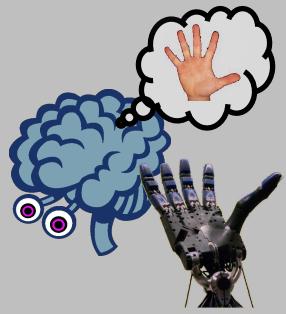
Lineare Regression



$$P = b + b_x x + b_y y + b_z z$$

Dabei ist **P** die **gemittelte spektrale Amplitude**, **x, y**, und **z** sind die kartesischen Handkoordinaten während des Haltens der Zielposition. Die **b**'s sind wiederum Koeffizienten, die mit **linearer Regression** zu bestimmen sind.





Lineare Regression

Allgemeine Form für einen Vektor \mathbf{x} :

$$y = b_0 + \sum_{n=1}^n b_n x_n + \epsilon$$

Matrix Form für Datenmatrix \mathbf{X} :

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \epsilon$$

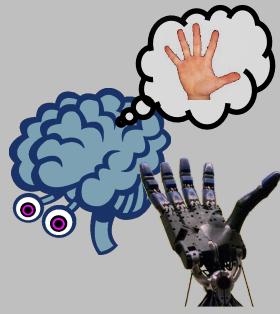
Diese hat folgende Dimensionalitäten:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{pmatrix}$$

\mathbf{X} ist die schon von der Klassifikation gewohnte Matrix aus n Featurevektoren mit p Features.

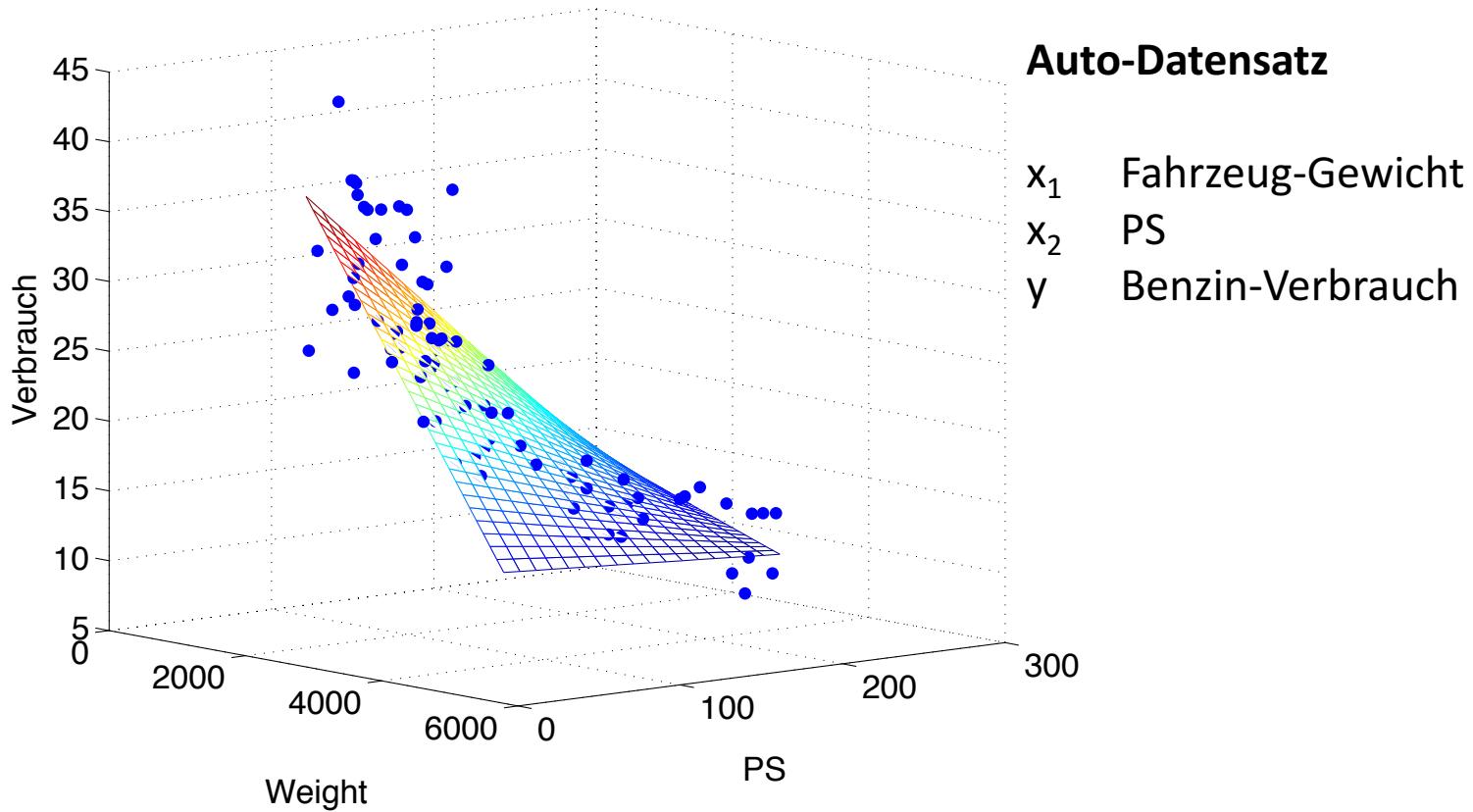
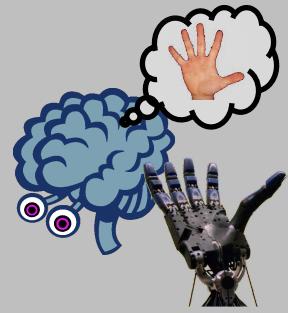
Für jedes Feature gibt es einen Koeffizienten, zusätzlich einen globalen Offset b_0 (auch y-intercept) genannt.

Lineare Regression

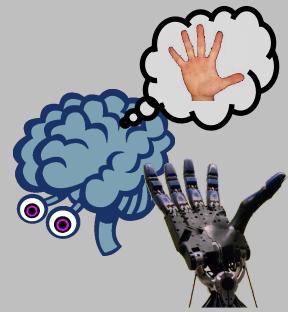


- Diese Darstellung stellt ein **Datenmodell** dar.
- Das Modell stellt eine explizite Beziehung zwischen den **Vorhersagevariablen X (predictor)** und den **Antwortvariablen Y (response)** dar.
- **Y** ist die **abhängige** und **X** die **unabhängige** Variable.
- Die Matrix **X** wird auch **Design-Matrix** genannt.
- Das Modell ist **linear in den Modell-Koeffizienten**, nicht notwendigerweise in der Basis-Funktion. Diese kann eine einfache Gerade sein, aber auch ein Polynom.
→ General Linear Models (GLM)

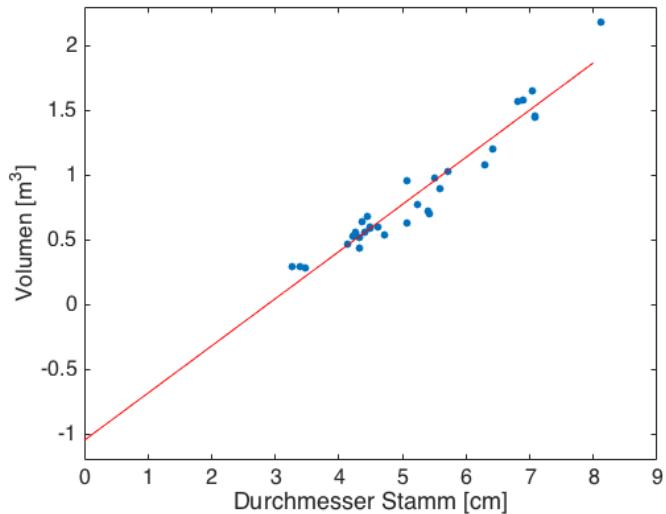
Einfache Beispiele



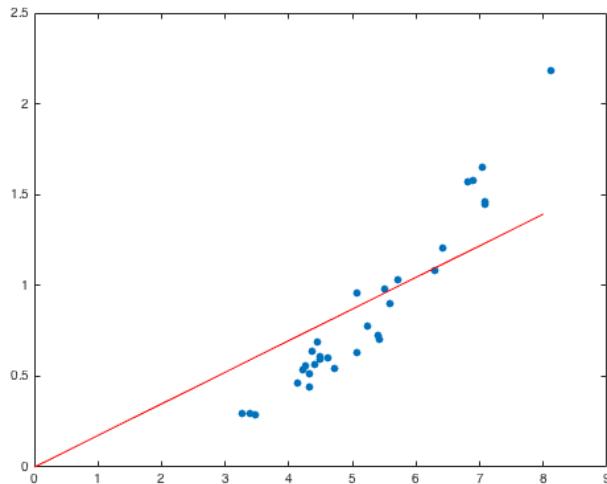
Die Bedeutung des Offset b_0



Mit Offset

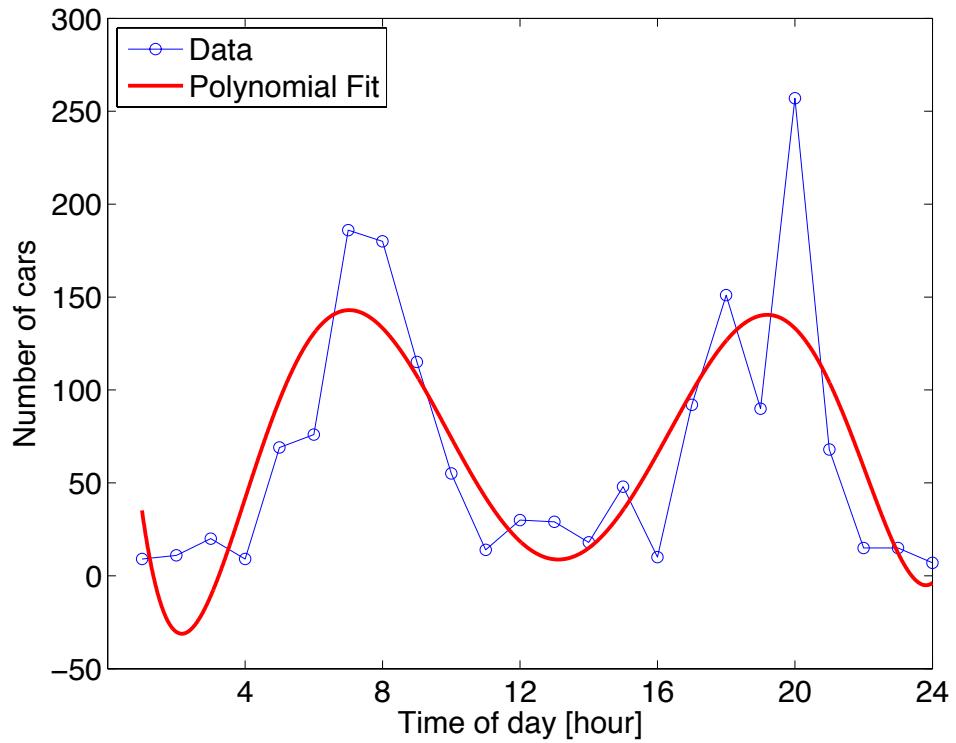
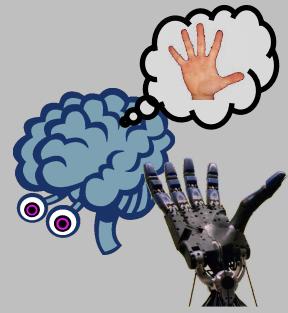


Ohne Offset



Beispieldatensatz Verhältnis des Stamm-Durchmessers zu Volumen der Baumkrone bei Kirschbäumen.

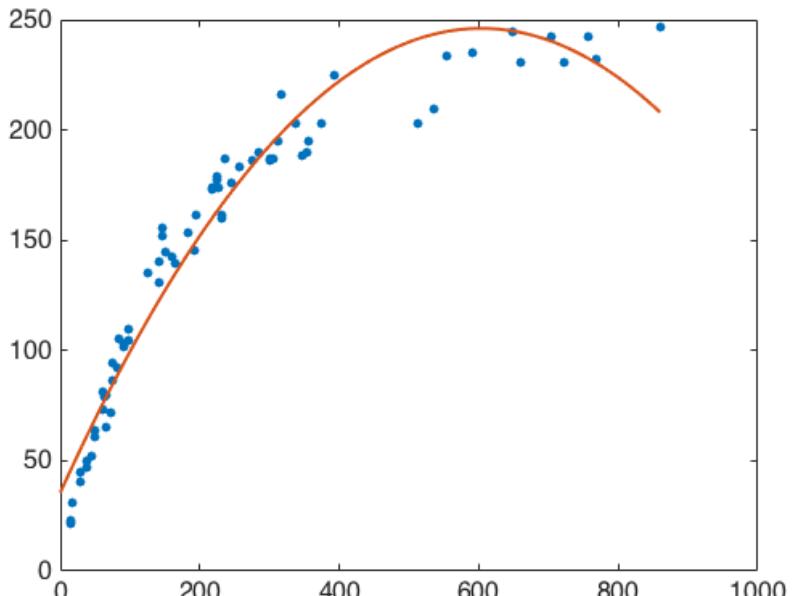
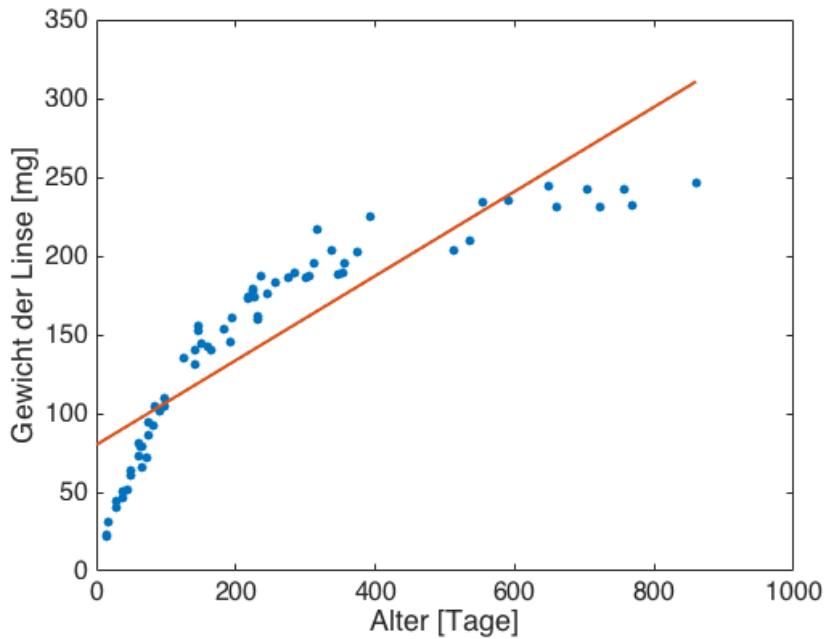
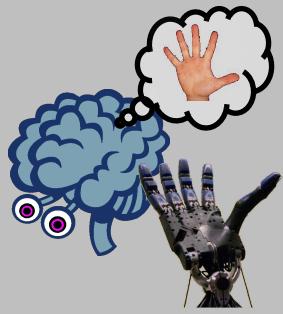
Einfache Beispiele



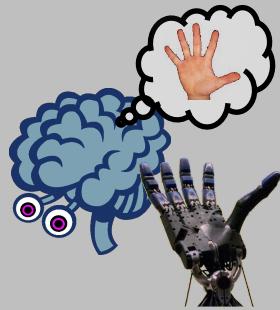
Verkehrsdatensatz

Fit eines Polynoms 6. Grades

Einfache Beispiele



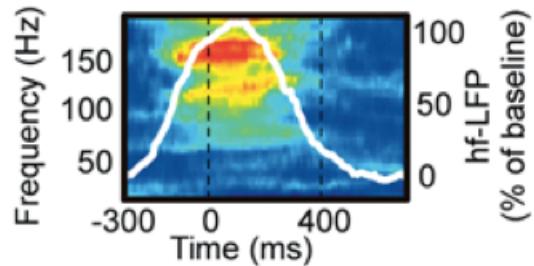
Verhältnis von Alter zu Gewicht der Linse (Auge) bei Kaninchen.



Zurück zum BMI Beispiel

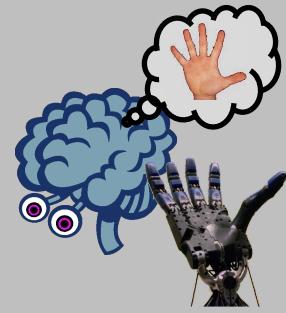
$$P = b + b_x x + b_y y + b_z z$$

Dabei ist **P** die **gemittelte spektrale Amplitude**, **x, y**, und **z** sind die kartesischen Handkoordinaten während des Haltens der Zielposition. Die **b**'s sind wiederum Koeffizienten, die mit **linearer Regression** zu bestimmen sind.



- Die Modellparameter **b** werden wie bei der Klassifikation aus einem “Trainingsdatensatz” bestimmt.
- Nach Bestimmung der Koeffizienten im “Training” können die relevanten Werte (hier: die Handpositionen) im Online-Lauf des Systems berechnet werden (“Umkehrung” des Modells).

Bestimmen der Modellparameter mit der *Least Squares* Methode



Schätzung der Koeffizienten:

$$\hat{\mathbf{b}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$



Moore-Penrose Pseudo-Inverse von \mathbf{X} (Verallgemeinerung der Matrixinversion für nicht-quadratische Matrizen)

Optimierungsproblem:

$$\hat{\mathbf{b}} = \underset{\mathbf{b}}{\operatorname{argmax}} \mathbf{S}(\mathbf{b})$$

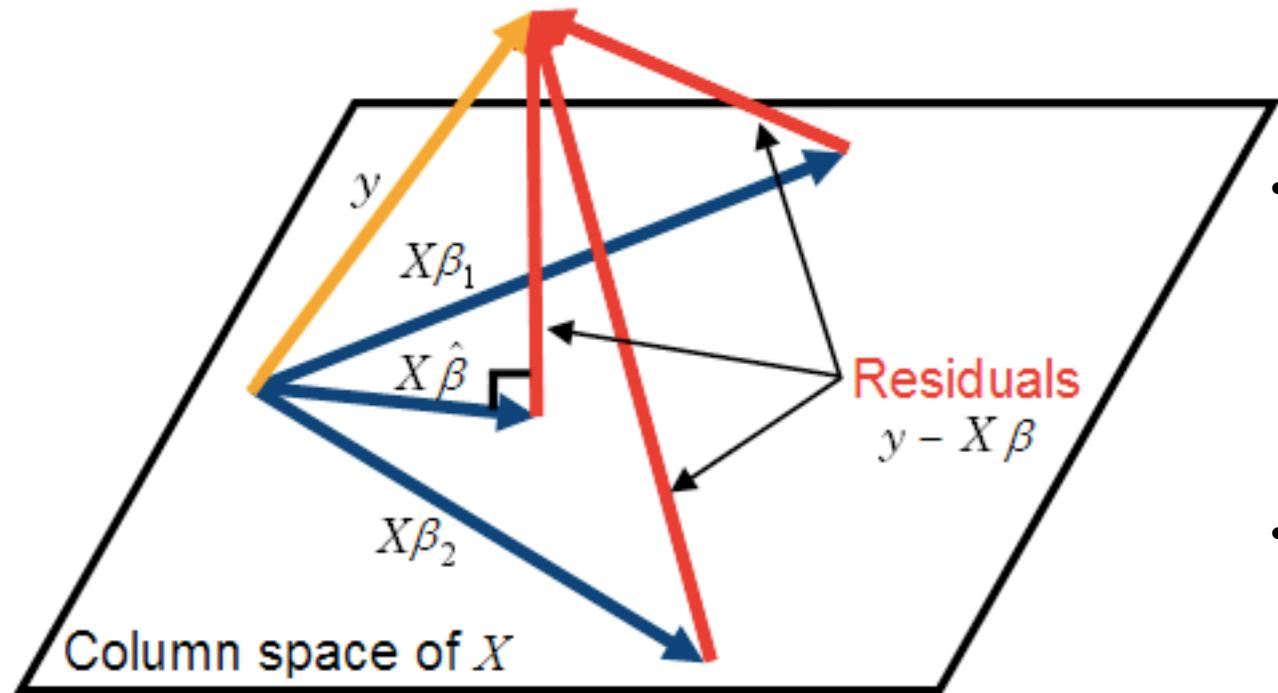
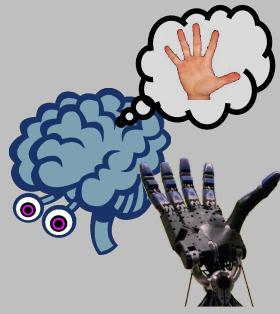
Fehlerfunktion S :

$$S(\mathbf{b}) = \|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2$$



Residuals

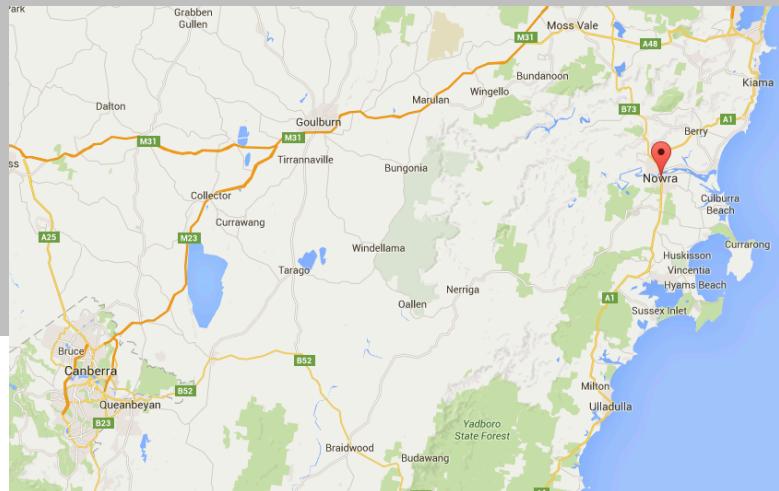
Least Squares – Geometrische Interpretation



- Wir suchen die Projektion des Datenvektors \mathbf{y} auf den von den Vektoren $\mathbf{X}\beta_1$ und $\mathbf{X}\beta_2$ aufgespannten Raum.
- Die Minimierung der Fehlerfunktion S entspricht der Minimierung der Länge der Residual-Vektoren.

Vorhersage mit trainiertem Modell

Beispiel Wetterdaten zweier Städte in Australien für das Jahr 1989



Datensatz

| Month | Day | Hour | Pressure | Wet Temp | Humidity | Direction | Speed | Rain | Dry Temp |
|-------|-----|------|----------|----------|----------|-----------|-------|------|----------|
| 01 | 01 | 02 | 947,2 | 15,1 | 14 | 00 | 0 | 4,0 | 14,6 |
| 01 | 01 | 05 | 947,1 | 14,0 | 13 | 04 | 2 | 5,0 | 13,5 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 12 | 31 | 23 | 948,0 | 22,0 | 9 | 00 | 0 | 0,0 | 15,5 |

$$\underline{X}$$

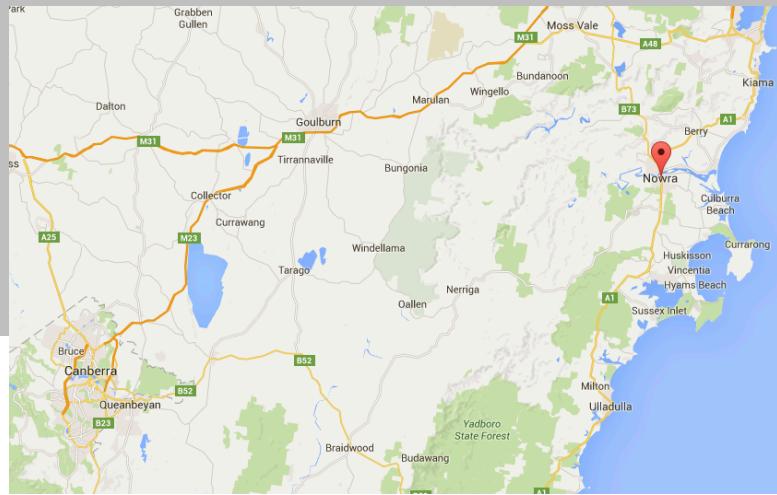
$$y$$

Trainingsdaten: Canberra

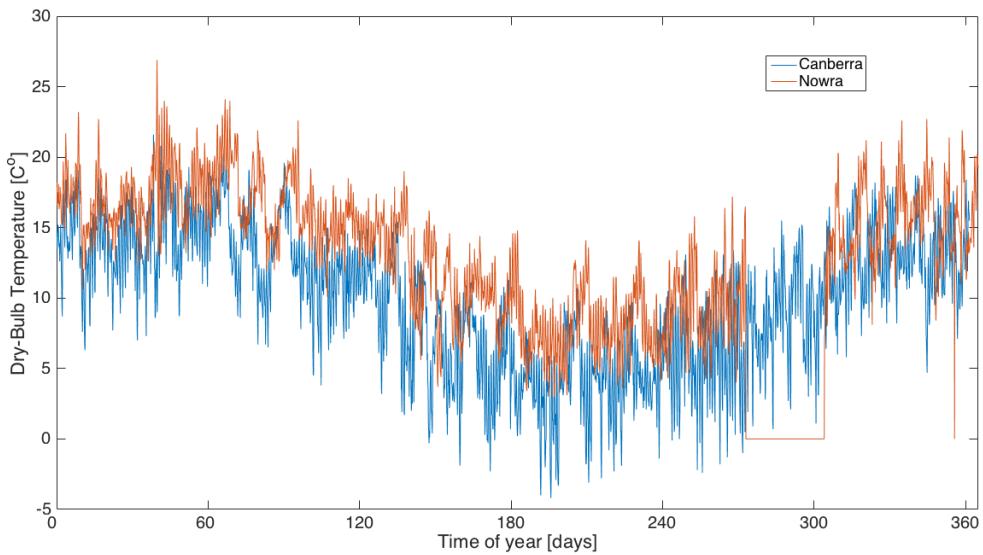
Testdaten: Nowra

Vorhersage mit trainiertem Modell

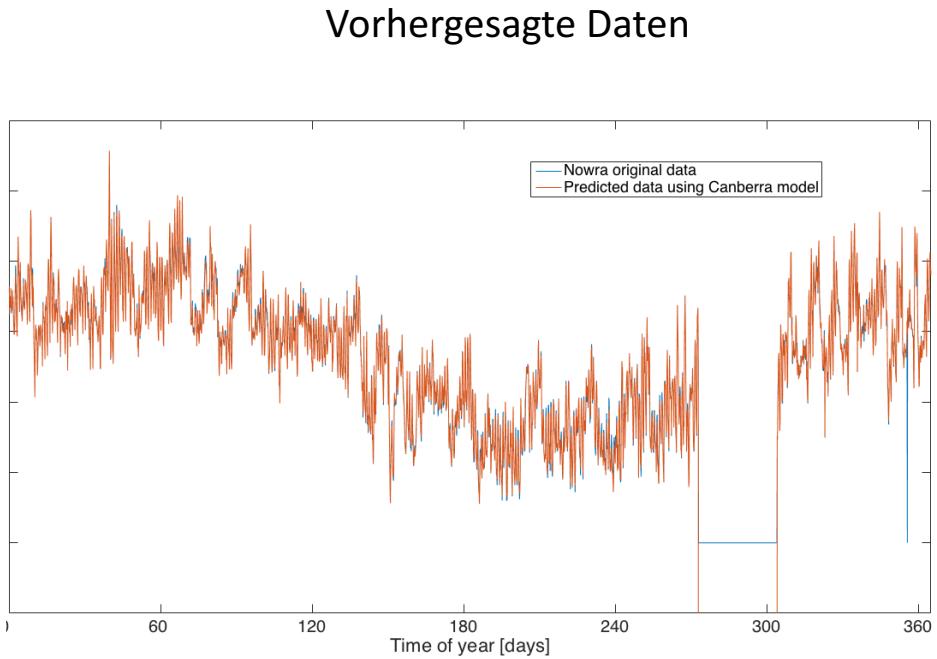
$$y_{pred} = b_0 + \sum_{i=1}^N \vec{x}_i b_i$$



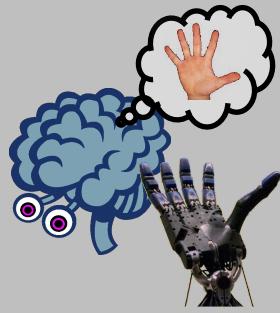
Originale Daten



Vorhergesagte Daten



Zusammenhang ANOVA <--> Regression



Beispieldatensatz: Klassifikationsergebnisse Datensatz Fixation-Related Potentials

1 Faktor (UV): Kanäle. 3 Faktorstufen: Nur EEG, Nur EOG, EEG+EOG. 10 Versuchspersonen.

ANOVA

| Vergleich der Mittelwerte je Faktor* | EEG | EOG | EEG+EOG |
|--------------------------------------|------|------|---------|
| | 0,88 | 0,68 | 0,82 |

F = 11,17 p = 0,0003 → statistisch signifikant

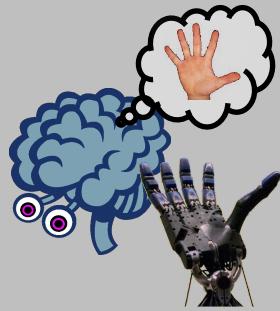
Regression

Gesucht wird ein Modell $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$, das den Zusammenhang zwischen den Messwerten y (Gewichtszunahme) und den Faktorstufen (x_1, x_2, x_3) erklärt. Die Faktorstufen werden hier auch Kategorien oder kategoriale Variablen genannt (im Gegensatz zu numerischen Variablen).

Dazu ist zunächst eine s.g. "Dummymodellierung" der Faktorstufen/ Kategorien notwendig.

*Maß: Fläche unter der ROC → AUC

Zusammenhang ANOVA <--> Regression



Die Daten liegen für die ANOVA in Tabellenform vor, d.h. die Variable y geht in der Tabelle auf und liegt nicht wie für die Regression benötigt, als Vektor vor.

Die Dummykodierung erzeugt eine neue Matrix, welche die Zugehörigkeit zu einer Faktorstufe binär darstellt.

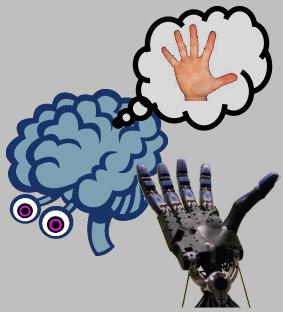
| ANOVA | | | Regression | | | | |
|----------|----------|----------|------------|---------|-------|----------------------------------|-------|
| A1 | A2 | A3 | | off-set | x_1 | x_2 | x_3 |
| y_{11} | y_{12} | y_{13} | y_{11} | 1 | 1 | 0 | 0 |
| y_{21} | y_{22} | y_{23} | y_{21} | 1 | 1 | 0 | 0 |
| 0,88 | 0,68 | 0,82 | y_{12} | 1 | 0 | 1 | 0 |
| | | | y_{22} | 1 | 0 | 1 | 0 |
| | | | y_{13} | 1 | 0 | 0 | 1 |
| | | | y_{23} | 1 | 0 | 0 | 1 |
| | | | | y | X | Referenzgruppe (Designmatrix) | |

Koeffizienten:

$$b_0 = 0,82 \quad [b_0 = a_3]$$

$$b_1 = 0,06 \quad [a_1 - b_0 = b_1]$$

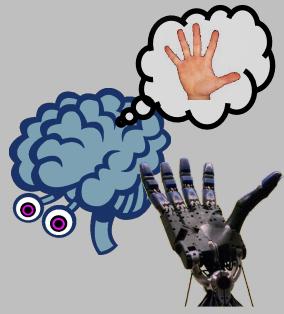
$$b_2 = -0,14 \quad [a_2 - b_0 = b_2]$$



Stationarität der Modelle

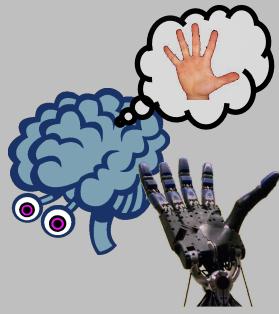
- In dem vorgestellten allgemeinen Modell sind die Koeffizienten b fix für alle Zeitpunkte der Messung (gegeben durch die Sample-Vektoren x_n).
- Dies entspricht der Annahme, dass die Parameter stationär sind, sich also nicht mit der Zeit ändern (das gilt im Übrigen ebenso für Klassifikator-Parameter).
- Diese Annahme ist nicht immer richtig, insbesondere sind physiologische Daten von ihrer Natur her immer instationär, d.h. ihr Eigenschaften variieren über die Zeit.
- Es gibt lineare Modelle, die solche Instationaritäten inkorporieren. Diese werden *Time-Varying Linear Regression* genannt. Die Koeffizienten des Modells werden dann abhängig von der Zeit: $b_i(t)$. Allerdings werden solche Modelle sehr selten verwendet.
- Statt dessen verwendet man **dynamische lineare Systeme**, wie z.B. den **Kalman-Filter**. Solche Modelle sind allerdings deutlich rechenintensiver als simple Regressionsmodelle.
- Lineare Regression ist **das** Standardverfahren zur Auswertung von Daten aus bildgebenden Verfahren. Dann wird es **Statistical Parametric Mapping** genannt.

Wiener Filter



- Der Wiener Filter wurde 1949 von Norbert Wiener publiziert.
- Er ist ebenfalls ein Verfahren der multi-dimensionalen Regression.
- Das Ziel ist die statistische Schätzung eines unbekannten Signals aus einem gegebenen, verrauschten Signal.
- Der Begriff „Filter“ stammt daher, dass eben dieses unbekannte Signal dadurch berechnet wird, indem das Rauschen entfernt wird.

Wiener Filter - Formulierung

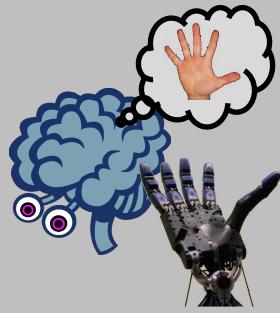


- Die Formulierung des Wiener Filters ist allgemein:

$$\mathbf{y}(t) = \mathbf{b} + \sum_{u=1}^n \mathbf{a}(u) \mathbf{x}(t-u) + \epsilon(t)$$

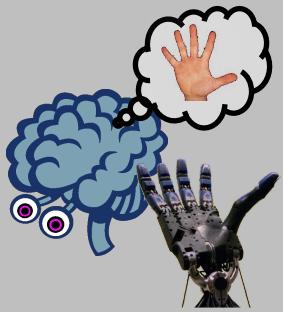
- Dabei ist allgemein
 - $\mathbf{y}(t)$ die Ausgabe des Filters zum Zeitpunkt t
 - \mathbf{b} eine Konstante (der s.g. y-Intercept),
 - $\mathbf{a}(u)$ wiederum die gesuchten Koeffizienten
 - $\mathbf{x}(t-u)$ der Wert des Eingangssignal zum Zeitpunkt $t - u$ („Historie“)
 - ϵ das additive Rauschen (bzw. der Rekonstruktionsfehler)
- Die Ausgabe $\mathbf{y}(t)$ ist im einfachsten Fall einfach die Schätzung des nächsten Samples gegeben eine Menge vergangener Samples.
- Die Filterordnung ist n , also die Anzahl der vergangenen Samples, die in die Schätzung der Ausgabe eingehen.

Wiener Filter - BMI



- Bei Verwendung im BMI ist die Ausgabe $\mathbf{y}(t)$ wie schon zuvor das gewünschte Steuersignal (z.B. x,y,z Koordinaten, Gelenkwinkel, Kraft,...)
- Das Modell ist stationär für n Samples (die Filterordnung).
- Bei Verwendung eines gleitenden Fensters der Größe n über ein kontinuierliches Signal erhält man somit ein Verfahren, das der Instationarität physiologischer Signale Rechnung trägt.
- Die Berechnung der gesuchten Koeffizienten $\mathbf{a}(u)$ erfolgt in Matrix-Form $\mathbf{Y} = \mathbf{X}\mathbf{A}$ und somit in der schon bekannten Form

$$\mathbf{A} = \text{inv}(\mathbf{X}^T \mathbf{X}) \mathbf{X}^T \mathbf{Y}$$



Wiener Filter - BMI

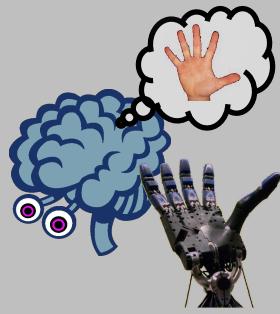
Die Matrix-Form $\mathbf{Y} = \mathbf{XA}$ hat bei einer mehrkanaligen Aufnahme die folgende Form

$$\begin{pmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_n) \end{pmatrix} =$$

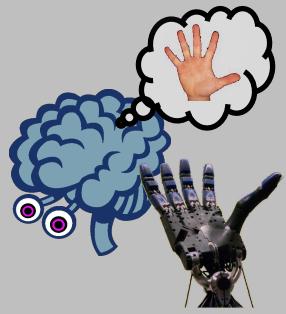
$$\begin{pmatrix} 1 & x^{c1}(t_1 - 1) & x^{c1}(t_1 - 2) & \dots & x^{c1}(t_1 - u) & x^{c2}(t_1 - 1) & \dots & x^{ck}(t_1 - u) \\ 1 & x^{c1}(t_2 - 1) & x^{c1}(t_2 - 2) & \dots & x^{c1}(t_2 - u) & x^{c2}(t_2 - 1) & \dots & x^{ck}(t_2 - u) \\ \vdots & \vdots \\ 1 & x^{c1}(t_n - 1) & x^{c1}(t_n - 2) & \dots & x^{c1}(t_n - u) & x^{c2}(t_n - 1) & \dots & x^{ck}(t_n - u) \end{pmatrix} * \begin{pmatrix} a_1(t_1 - 1) & a_1(t_1 - 2) & \dots & a_1(t_1 - u) \\ a_2(t_2 - 1) & a_2(t_2 - 2) & \dots & a_2(t_2 - u) \\ \vdots & \vdots & \vdots & \vdots \\ a_n(t_n - 1) & a_n(t_n - 2) & \dots & a_n(t_n - u) \end{pmatrix}$$

- $y(t_i)$ kinematischer Parameter zum Zeitpunkt t_i
- $x_{cj}(t_u)$ Signal von Kanal j zum Zeitpunkt t_u
- $ak(t_u)$ gesuchte Koeffizienten

Kalman Filter



- Der Kalman Filter wurde 1960 von R.E. Kalman publiziert.
- Er stellt eine rekursive Schätzung des Zustands eines Prozesses dar.
- Der Algorithmus minimiert den mittleren quadratischen Fehler.
- Der Kalman Filter erlaubt die Schätzung von vergangenen, aktuellen oder zukünftigen Zuständen eines Prozesses.
- Die Eigenschaften des zu modellierenden Systems müssen nicht genau bekannt sein.



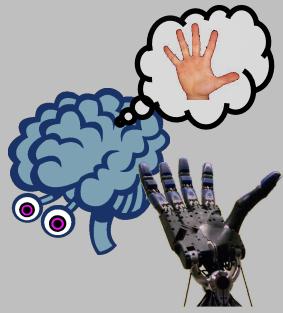
Kalman Filter - Formulierung

- Der Kalman-Filter geht von der Signaldarstellung im Zustandsraum aus und führt dort zu einem Filter, das den Prozess nachbildet.
- Der **Zustand** wird durch einen Vektorprozess $\mathbf{x}(n)$ beschrieben, einem Input-Steuervektor \mathbf{u} , und einem additiven Rauschen (unkorreliert, mittelwertfrei) **Rauschquellen** $\mathbf{w}(n)$ erregt wird.
- Die **Messsignale** $\mathbf{y}(n)$ sind zusätzlich von unkorrelierten, mittelwertfreien Rauschquellen $\mathbf{z}(n)$ überlagert.
- Die Rauschquellen des Vektorprozesses und der Messsignale sind unabhängig.
- Das allgemeine Modell hat die Gestalt:

$$\mathbf{x}(n + 1) = D(n) \mathbf{x}(n) + B(n) \mathbf{u}(n) + \mathbf{w}(n) \quad \text{"process equation" (z.B. 2d Curser Pos.)}$$

$$\mathbf{y}(n) = C(n) \mathbf{x}(n) + \mathbf{z}(n) \quad \text{"observation equation" (Neuronal activity)}$$

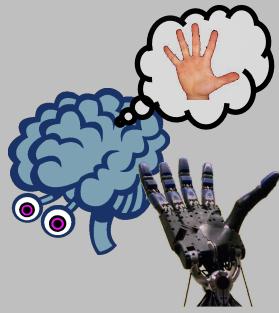
Kalman Filter - Formulierung



- Die Ordnung M des Zustandsvektors $x(n)$ ist durch die Ordnung des Systems gegeben.
- $D(n)$ ist die $(M \times M)$ – Systemmatrix.
- Die L Eingangsgrößen sind im Steuervektor $u(n)$ und die K Ausgangsgrößen im Meßvektor $y(n)$ zusammengefaßt.
- $B(n)$ ist die $(M \times L)$ – Eingangsmatrix und $C(n)$ die $(K \times M)$ – Ausgangsmatrix.
- Die Kovarianzmatrizen berechnen sich nach

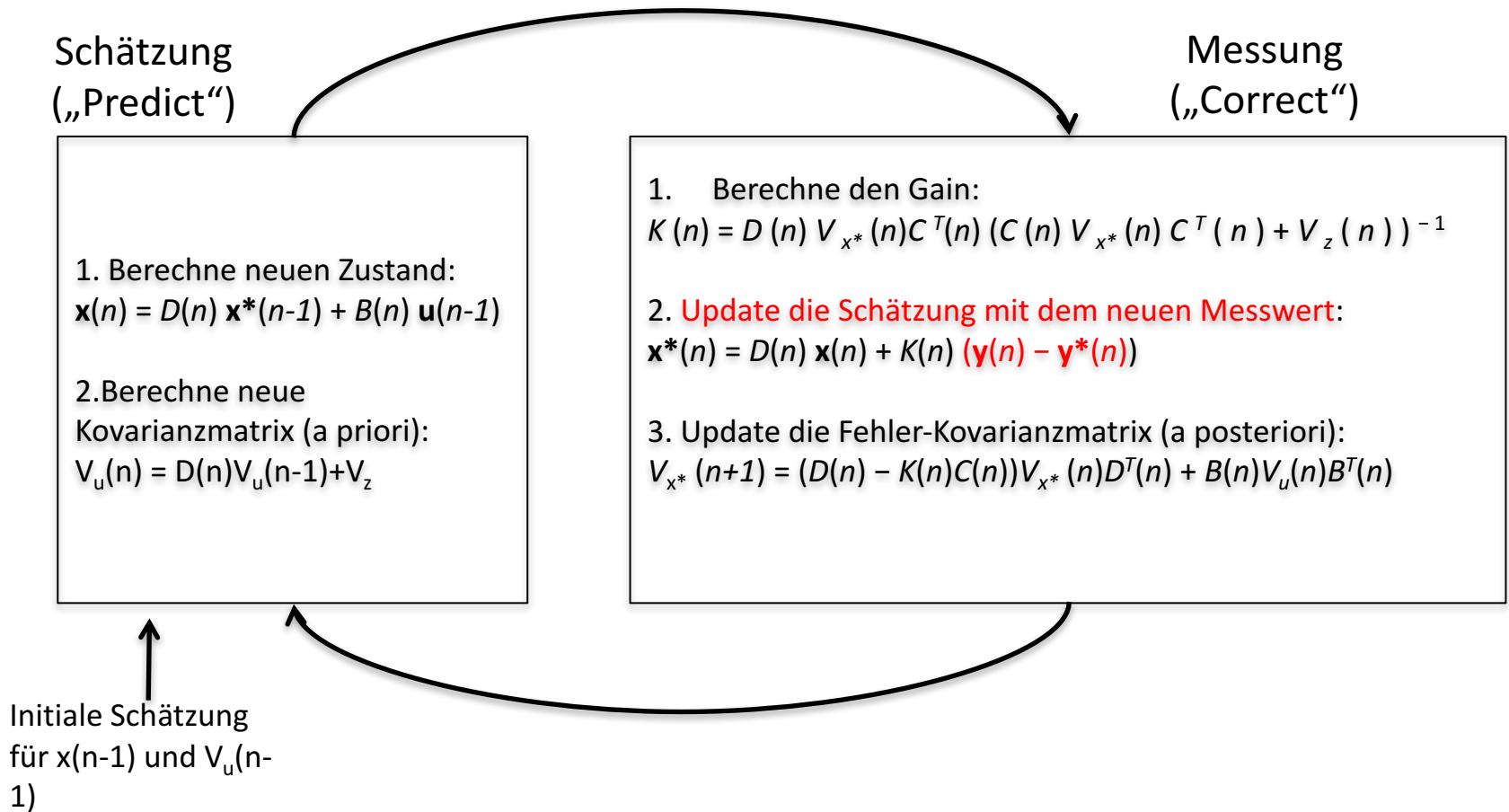
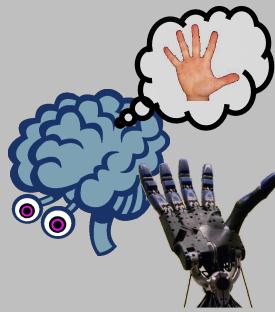
$$V_u(n) = (\mathbf{E}[u_i(n)u_j(n)])$$
$$V_z(n) = (\mathbf{E}[z_i(n)z_j(n)])$$

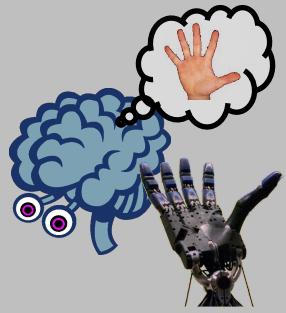
Kalman Filter - Formulierung



- Der neue Zustand $\mathbf{x}(n+1)$ errechnet sich aus dem alten Zustand $\mathbf{x}(n)$ und den neuen Meßwerten $\mathbf{y}(n)$ nach dem Gütekriterium des minimalen quadratischen Schätzfehlers (“least mean squares”).
- Schätzungen $\mathbf{x}^*(n)$ für den Zustand und $\mathbf{y}^*(n)$ für die Messgrößen rekursiv aus
$$\mathbf{x}^*(n + 1) = D(n) \mathbf{x}^*(n) + K(n) (\mathbf{y}(n) - \mathbf{y}^*(n))$$
$$\mathbf{y}^*(n) = C(n) \mathbf{x}^*(n)$$
- Die **Gewichtsmatrix oder Gewinnfunktion (Gain)** des Kalman-Filters $K(n)$ ergibt sich aus
$$K(n) = D(n) V_{x^*}(n) C^T(n) (C(n) V_{x^*}(n) C^T(n) + V_z(n))^{-1}$$
- wobei die Kovarianzmatrix $V_{x^*}(n)$ rekursiv aus
$$V_{x^*}(n + 1) = (D(n) - K(n)C(n))V_{x^*}(n)D^T(n) + B(n)V_u(n)B^T(n)$$

Kalman Filter - Schema





Anwendungsbeispiel

Dekodierung von Bewegungsparametern (Geschwindigkeit, Beschleunigung) in 3D für BMI-basierte Cursorkontrolle bei Affen (invasive Elektroden)*

Neuronales Signal: Diskretisierte Feuerrate (Spikes), mit $dT = 90 \text{ ms}$)

- Vereinfachende Annahmen:

- Der Steuervektor u wird als unbekannt angesehen und auf Null gesetzt.
- Die Systemmatrix D , die Ausgangsmatrix C , sowie die Rauschvektoren w und z werden als stationär angenommen.

Damit vereinfachen sich die Gleichungen zu:

$$\mathbf{x}(n+1) = D(n) \mathbf{x}(n) + B(n) \mathbf{u}(n) + \mathbf{w}(n) \rightarrow \mathbf{x}(n+1) = D\mathbf{x}(n) + \mathbf{w} \quad [\text{Modell des Curser Zustands}]$$

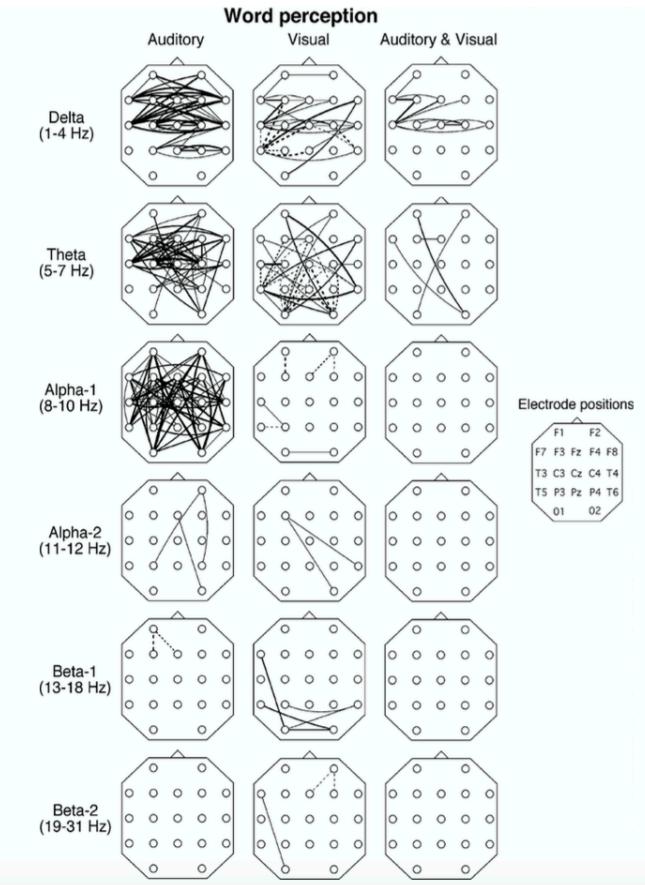
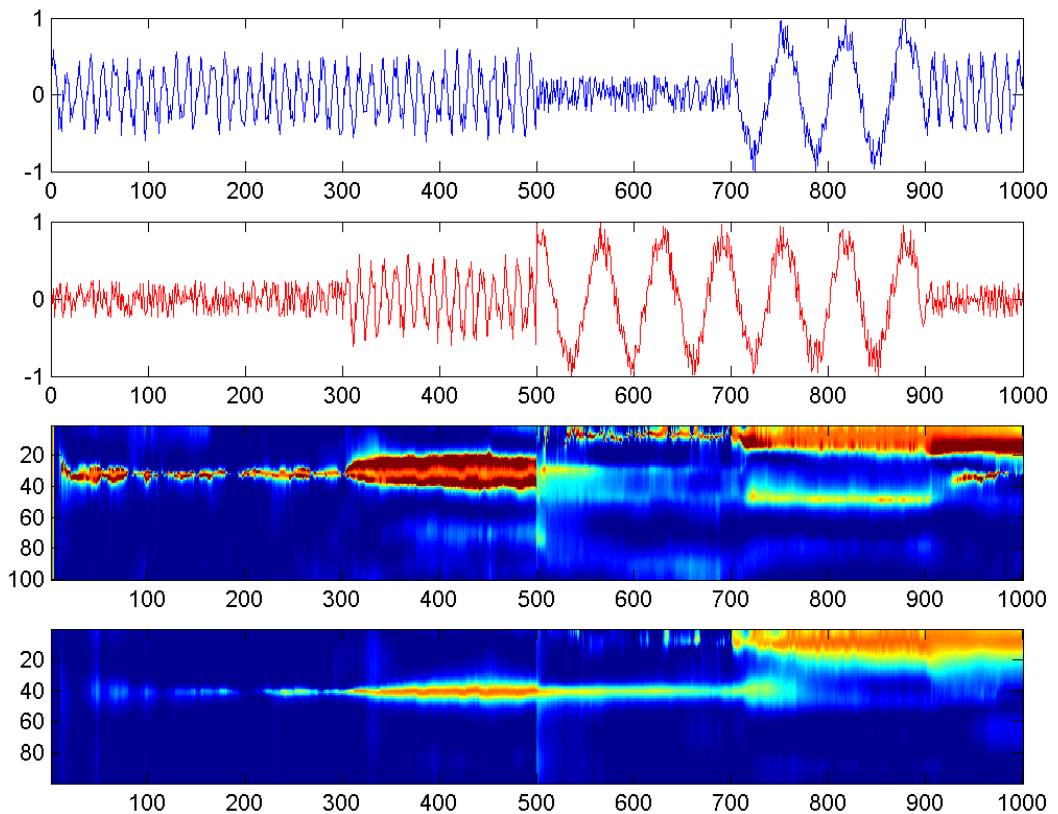
$$\mathbf{y}(n) = C(n) \mathbf{x}(n) + \mathbf{z}(n) \rightarrow \mathbf{y}(n) = C\mathbf{x}(n) + \mathbf{z} \quad [\text{Modell der neuronalen Aktivität}]$$

Die Updateschritte werden wie zuvor berechnet.

*Hauschild et al, PNAS, 2012

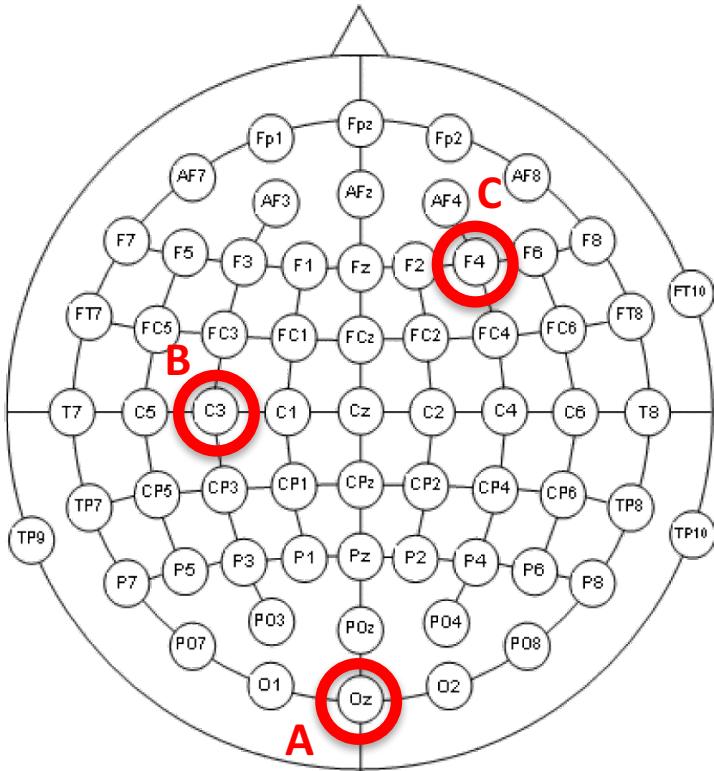
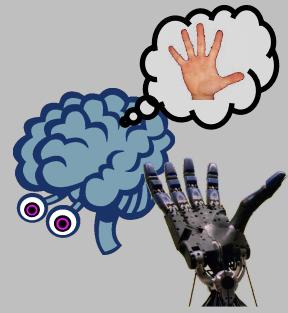
Recap: Coherence and Phase Locking

„Toy Example“

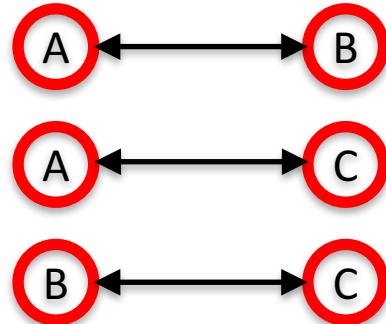


Studie: Vergleich der Kohärenzen bei der Perzeption konkreter und abstrakter Wörter.

Undirected vs. Directed Dependencies

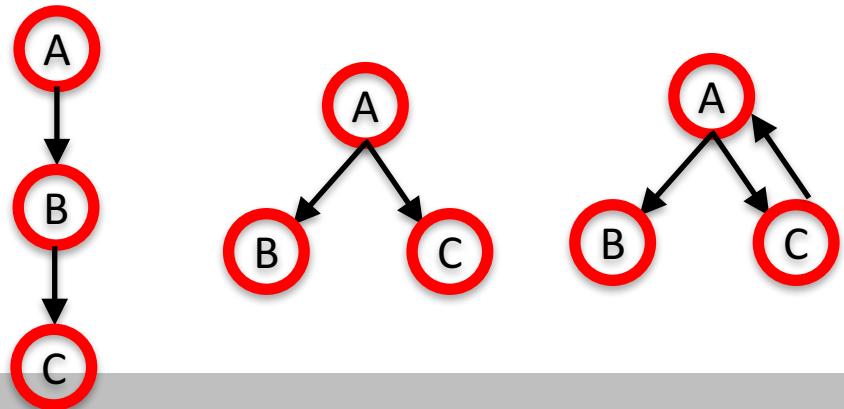


Undirected: Coherence and Phase-Locking

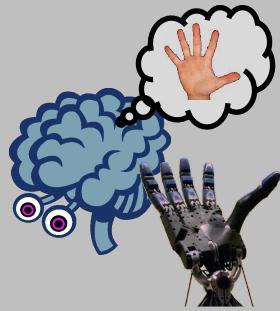


Channels
are only
analyzed
pairwise!

Directed: Granger Causality

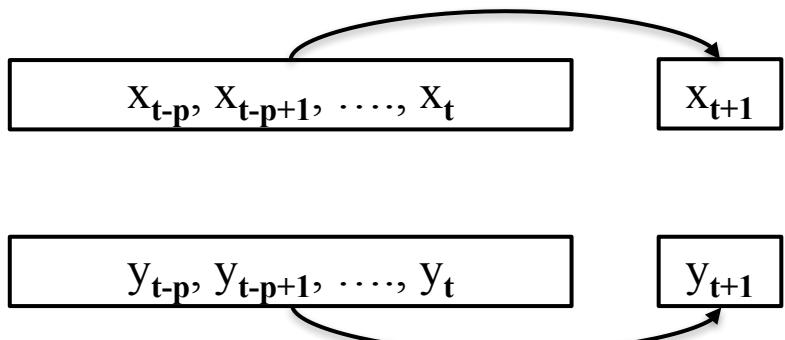


Granger Causality

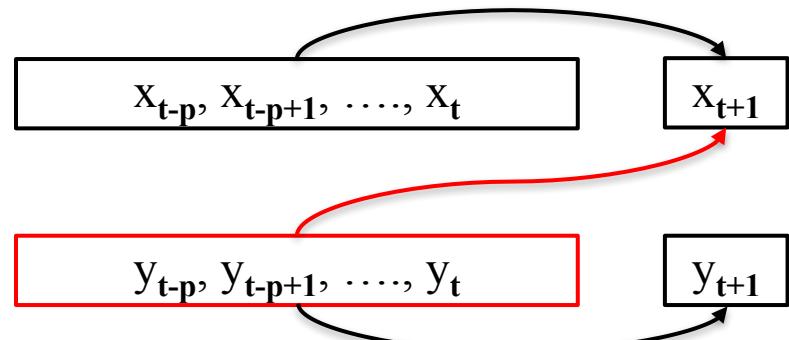


The general idea

Suppose that we have two variables X and Y, and we try to predict X_{t+1} using only past terms of X. We also try to predict X_{t+1} using past terms of both X and Y. If the second prediction is significantly more successful, then the past of Y contains information useful for predicting X_{t+1} that is not in the past of X. In this case, **Y is said to G-cause X**.

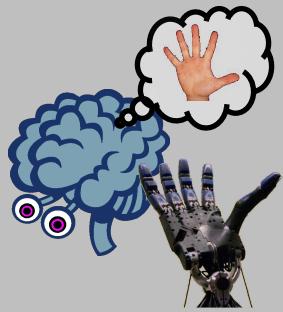


No G-causality



Y G-causes X

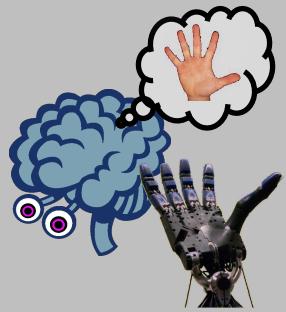
Formalization: VAR-Model with causal relations



- Uni-variate auto-regression model
$$x(t) = \sum_{j=1}^p a(t)x(t-j) + \epsilon$$
 - Prediction model:
→ Value $x(t)$ of a time series is predicted from its own history of length p , in other words, **depends** only on its own history
- Now consider a bi-variate auto-regression model with dependencies

$$X_1(t) = \sum_{j=1}^p A_{11}(j)X_1(t-j) + \sum_{j=1}^p A_{12}(j)X_2(t-j) + E_1(t)$$

$$X_2(t) = \sum_{j=1}^p A_{21}(j)X_1(t-j) + \sum_{j=1}^p A_{22}(j)X_2(t-j) + E_2(t)$$



MVAR Model with causal relations

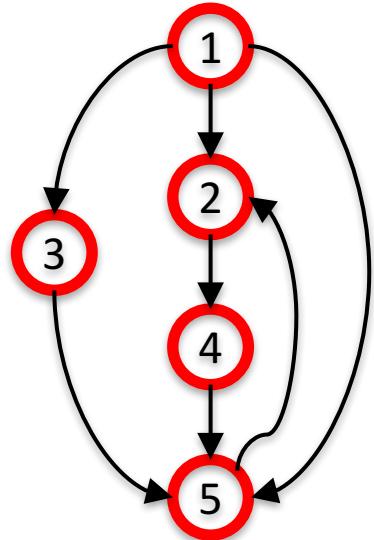
- Multi-variate AR model with dependencies
- Example

$$\mathbf{X}(t) = \sum_{i=1}^p \mathbf{A}(i)\mathbf{X}(t-1) + \mathbf{E}(t)$$

```

X(1,n) = 1.58*X(1,n-1) - 0.81*X(1,n-2) + W(1,n);
X(2,n) = 0.9*X(1,n-1) - 0.01*X(2,n-2) - 0.6*X(5,n-3) + W(2,n);
X(3,n) = 0.3*X(2,n-1) + 0.8*X(3,n-1) + 0.3*X(2,n-2) - 0.25*X(3,n-2) + 0.3*X(2,n-3) + W(3,n);
X(4,n) = 0.9*X(4,n-2) - 0.6*X(2,n-1) + 0.3*X(2,n-3) + W(4,n);
X(5,n) = -0.3*X(4,n-1) + 0.9*X(1,n-2) - 0.3*X(4,n-2) + 0.6*X(3,n-3) + W(5,n);

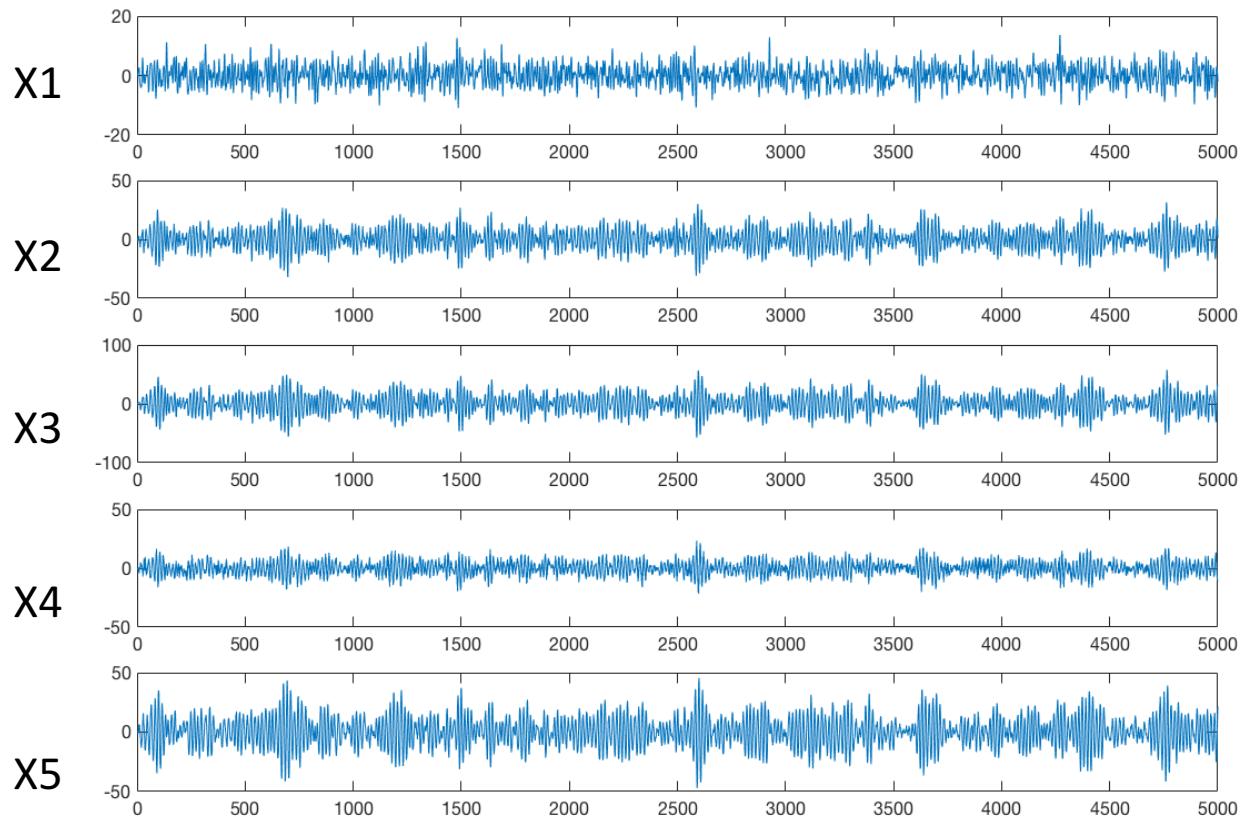
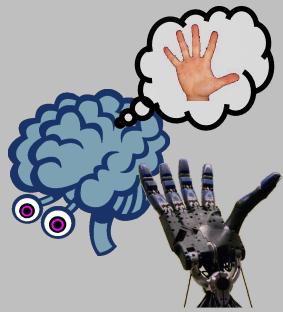
```



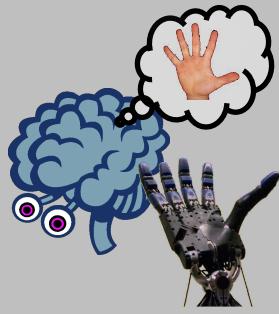
→ Complex pattern of directed interactions

For neural data, the underlying model and its causal dependencies are unknown.

MVAR Model with causal relations



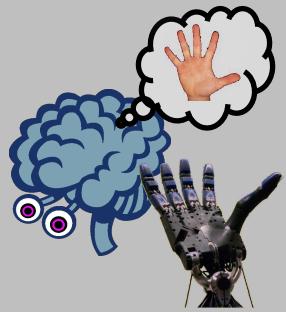
Causal Relations in the Spectral Domain



- Frequency analysis based on an MVAR model is a *parametric* technique.
- That is, because the *AR coefficients will be Fourier transformed*, instead of the plain data.
- In the bi-variate case (for simplicity, it's straight forward to extend to MVAR case), the Fourier transform of the bi-variate model is:

$$\begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} E_1(f) \\ E_2(f) \end{pmatrix} \quad \text{with}$$

$$A_{lm}(f) = \delta_{lm} - \sum_{j=1}^p A_{lm}(j) e^{-i2\pi f j} \quad \text{where} \quad \delta_{lm} = \begin{cases} 1 & \text{if } l = m \\ 0 & \text{if } l \neq m \end{cases}$$



G-Causality from Transfer Matrix

- Rewriting the Fourier transform of the model results in

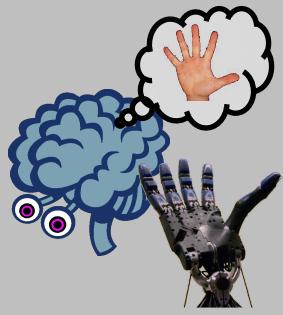
$$\begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{pmatrix} \begin{pmatrix} E_1(f) \\ E_2(f) \end{pmatrix} \quad \text{where}$$

$$\begin{pmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{pmatrix} = \begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix}^{-1}$$

- The Granger Causality in the spectral domain **from channel j to channel i** is now defined by

$$I_{j \rightarrow i}^2 = |H_{ij}(f)|^2 = \frac{|A_{ij}(f)|^2}{|\mathbf{A}(f)|^2}$$

Application



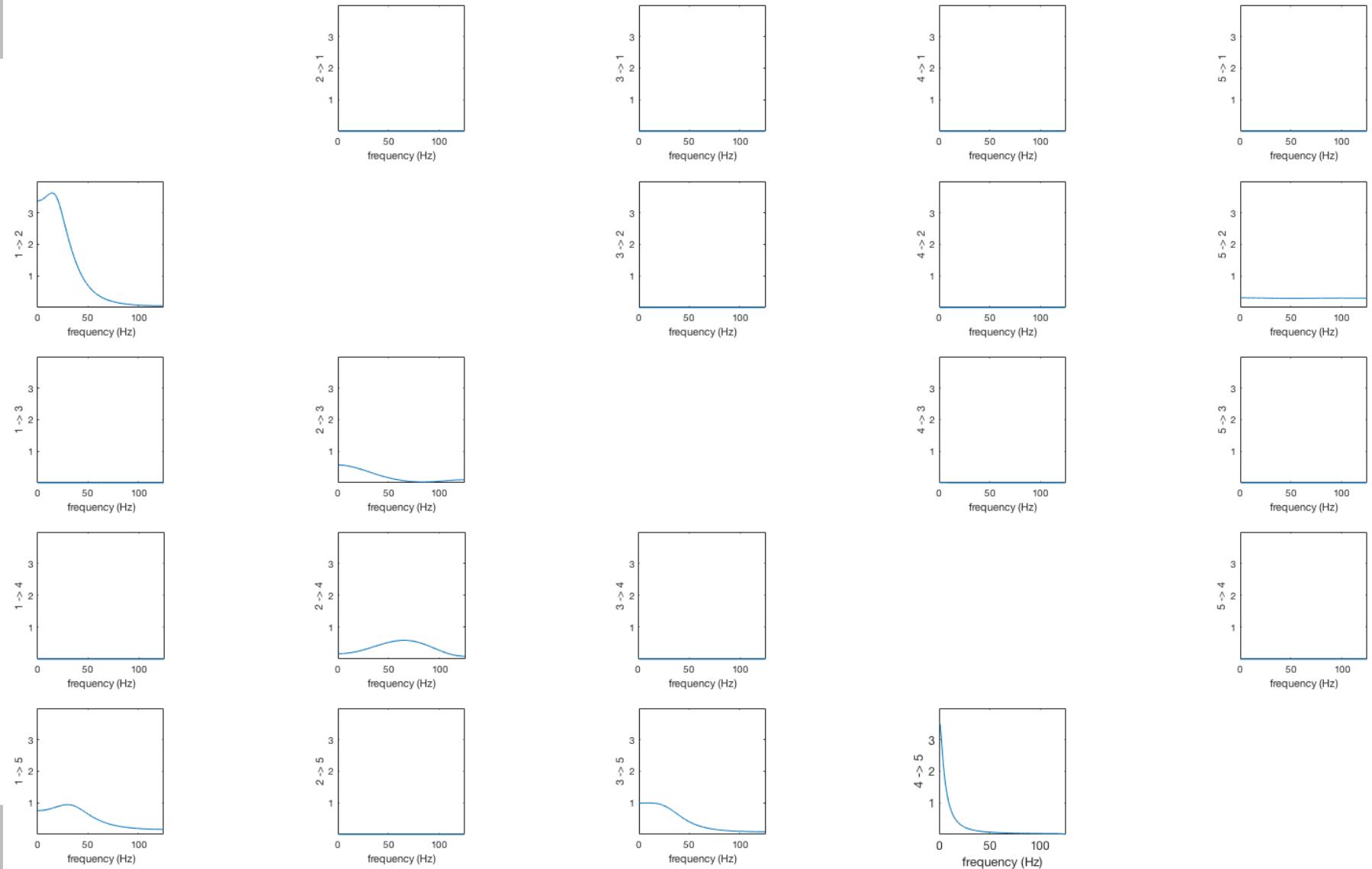
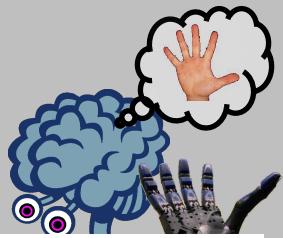
- The transfer matrix \mathbf{H} contains the auto-spectral components on the diagonal, and the cross-spectral components on the off-diagonal elements.
- Granger Causality is typically expressed either as
 - Direct Transfer Function
 - Partial Directed Coherence

$$\text{DTF}_{ij}(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^k |H_{im}(f)|^2}$$

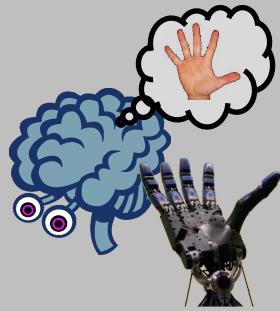
$$\text{PDC}_{ij}(f) = \frac{H_{ij}(f)}{\sqrt{h_j(f)^* \cdot h_j(f)}}$$

* complex conjugate

MVAR Example Model: Results



ARMA im instationären 1-D Fall



$$Y_n + \sum_{k=1}^p a_k(n) Y_{n-k} = \varepsilon_n - \sum_{j=1}^q b_j(n) \varepsilon_{n-j}$$

ARMA(p, q) – Modell mit variierbaren Parametern

Y_n Meßreihe

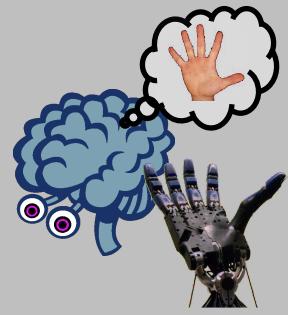
$a_k(n)$ autoregressiver Parameter

ε_n Rauschen

$b_j(n)$ moving average Parameter
(gleitender Mittelwert)

eindimensionales Signal

$$\{x_n\}_{n=1,2,\dots}$$



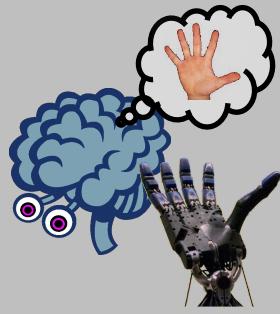
mit veränderlichen Struktureigenschaften

$$e_0 = 0$$

$$e_n = x_n + \sum_{k=1}^p \hat{a}_k(n-1)x_{n-k} + \sum_{j=1}^q \hat{b}_j(n-1)e_{n-j}, \quad n \geq 1$$

Vorhersagefehler zum Zeitpunkt n

Adaption der autoregressiven Parameter



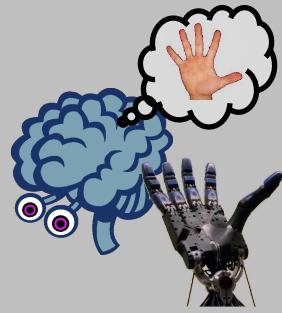
$$\hat{a}_k(n) = 0 , \quad n \leq k , \quad k=1, \dots, p$$

$$\hat{a}_k(n) = \begin{cases} d = \hat{a}_k(n-1) - c_n e_n x_{n-k} , & \text{falls } |d| < \binom{p}{k} , \quad n > k \\ \binom{p}{k} sgn(d) , & \text{sonst} \end{cases}$$



Abbruchbedingung, soll Parameterdrift verhindern

Adaption der Mittelwert Parameter

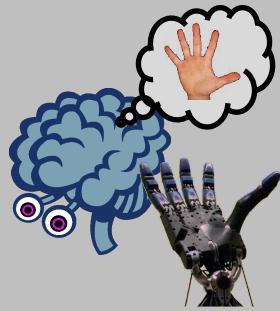


$$\hat{b}_j(n) = 0 , \quad n \leq j , \quad k = 1, \dots, q$$

$$\hat{b}_j(n) = \begin{cases} d = \hat{b}_j(n-1) - c_n e_n e_{n-j} , \\ \quad \quad \quad \text{falls } \sum_{l=1}^{j-1} |\hat{b}_l(n)| + |d| + \sum_{l=j+1}^q |\hat{b}_l(n-1)| < 1 \\ \hat{b}_j(n-1) , \quad \quad \text{sonst} \end{cases}$$

↑
Abbruchbedingung

Adaptionsgeschwindigkeit



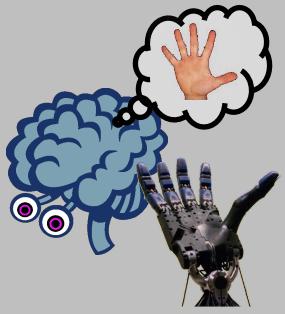
Konstruiert werden **Schätzfunktionenfolgen** der Modellparameter

Anpassung der autoregressiven Parameter schwerfälliger als der MA-Parameter

--> alleinige Korrektur der AR durch Vorhersagefehler

--> Wahl der Modellordnung: $p \gg q$

Vorhersagefehler



Vorhersagefehler enthält sogenannten Überschußfehler
(excess mean square error)

Minimierung: Wahl der Folge

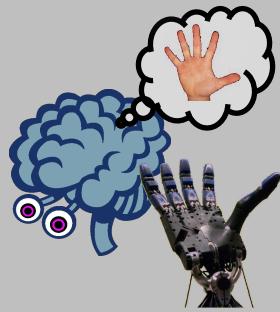
$$c_n = \frac{f}{1 + \sigma_x^2(n)}$$

konstanter positiver Faktor mit

$$0 < f < \frac{2}{p+q}$$

$\sigma_x^2(n)$ Varianz der Meßreihe zum Zeitpunkt n

Mehrdimensionales Modell



$$\mathbf{Y}_n + \sum_{k=1}^p A^k(n) \mathbf{Y}_{n-k} = \mathbf{Z}_n - B^j(n) \mathbf{Z}_{n-j}$$

d – dimensionales ARMA(p, q) - Modell

\mathbf{Y}_n Meßreihen

$A^k(n)$

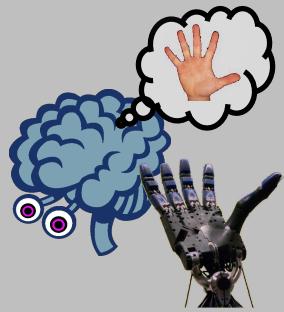
d x d – Matrix der AR-
Parameter

\mathbf{Z}_n Rauschen

$B^j(n)$

d x d – Matrix der MA-
Parameter

2-D Messsignal (2 Kanäle)



zweidimensionales Meßsignal

$$\{x_n = [x_n^1, x_n^2]\}_{n=0,1,\dots}$$

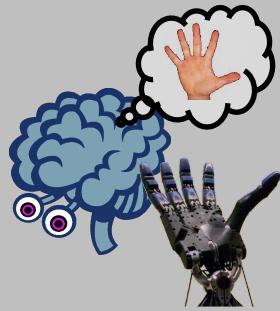
$$e_0 = 0$$

$$e_n = x_n + \sum_{k=1}^p \hat{A}^k(n-1)x_{n-k} + \sum_{j=1}^q \hat{B}^j(n-1)e_{n-j}$$

$$A^k = \begin{pmatrix} a_{11}^k & a_{12}^k \\ a_{21}^k & a_{22}^k \end{pmatrix}$$

$$B^j = \begin{pmatrix} b_{11}^j & b_{12}^j \\ b_{21}^j & b_{22}^j \end{pmatrix}$$

Adaption der Parameter



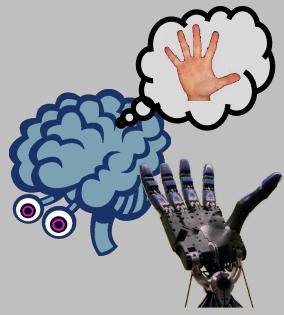
$$\hat{A}^k = 0 , \quad n \leq k , \quad k = 1, \dots, p$$

$$\hat{B}^j = 0 , \quad n \leq j , \quad j = 1, \dots, q$$

$$\hat{A}^k(n) = \hat{A}^k(n-1) - c_n e_n x_{n-k} , \quad n > k$$

$$\hat{B}^j(n) = \hat{B}^k(n-1) - c_n e_n e_{n-j} , \quad n > j$$

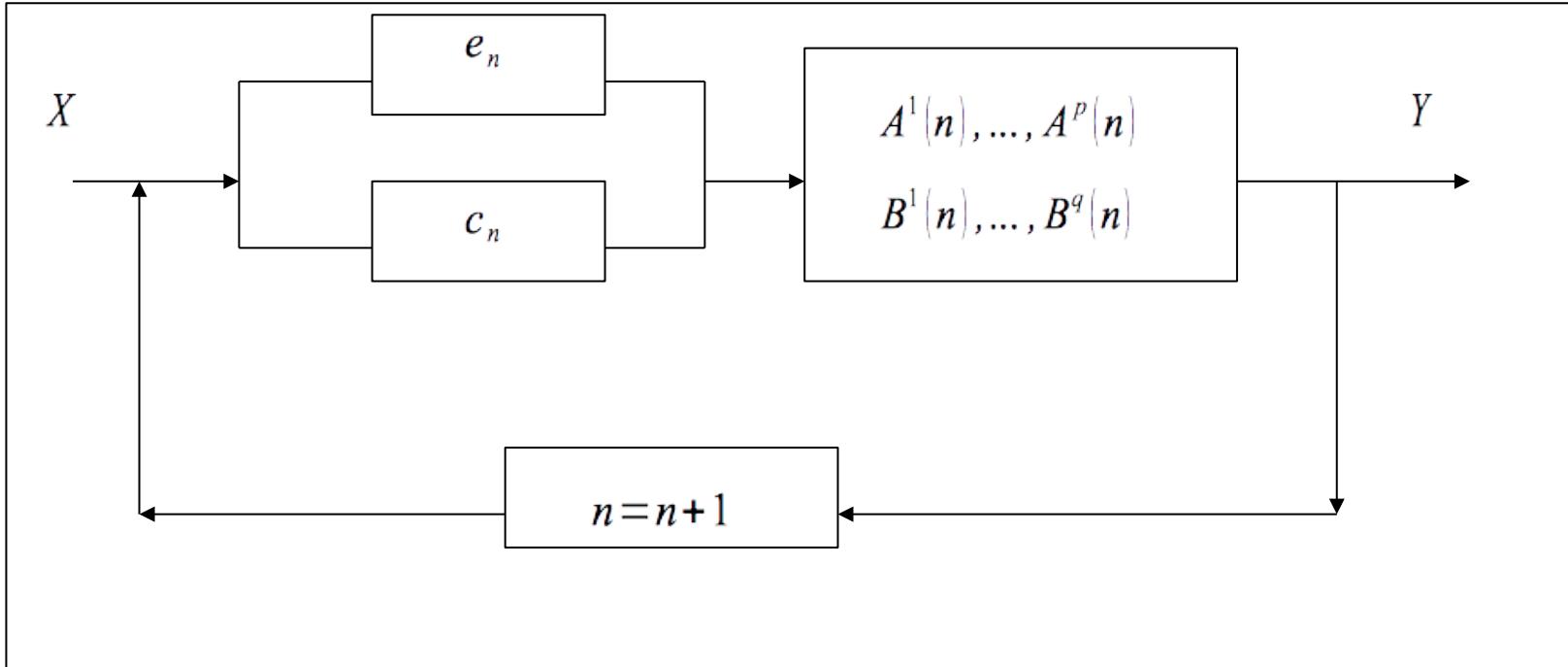
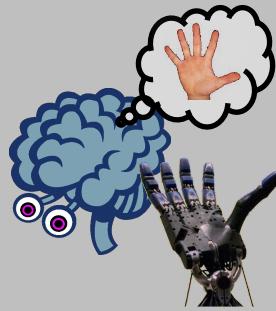
Adaption der Parameter



$$c_n = \frac{f}{1 + \hat{\sigma}_1^2(n) + \hat{\sigma}_2^2(n)}$$
 mit

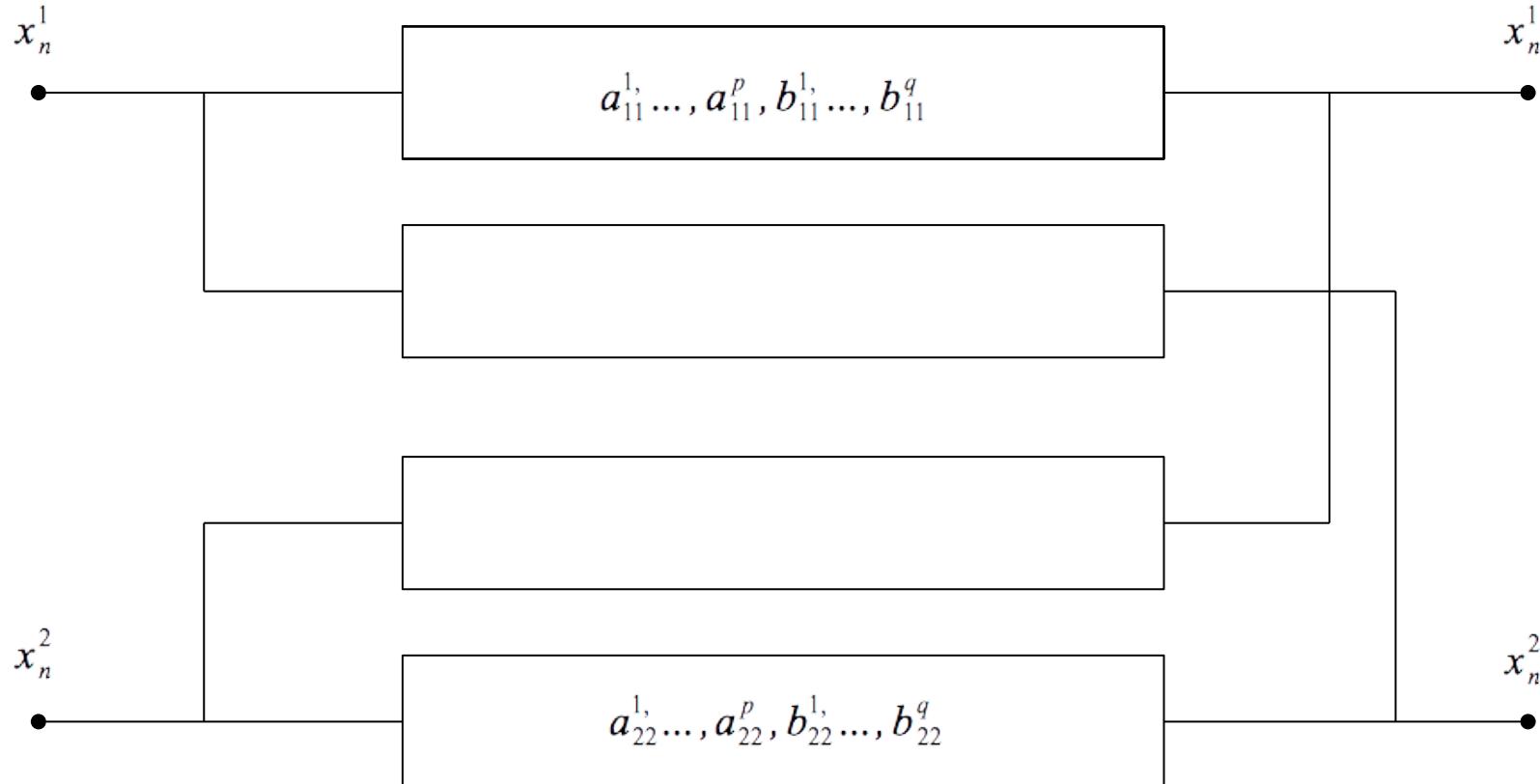
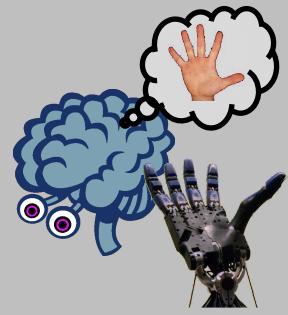
$$0 < f < \frac{2}{(p+q) \cdot 2}$$

Berechnungsschema

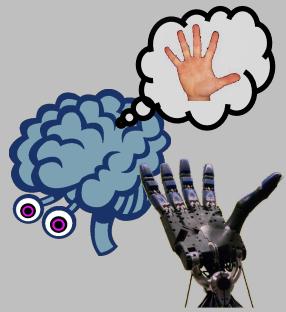


Zunächst Korrektur der Adoptionsvariablen c_n und des Vorhersagefehlers e_n , danach Aktualisierung der Parametermatrizen selbst

Berechnung des Abhängigkeitsgrades zweier Komponenten



--> Abhängigkeitsgrad errechnet sich aus der Summe der Kreuzparameter

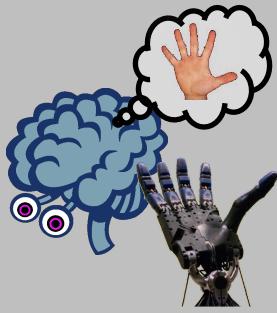


$$D(n) = \frac{\sum_{k=1}^p \lfloor (a_{12}^k)^2 + (a_{21}^k)^2 \rfloor + \sum_{j=1}^q \lfloor (b_{12}^j)^2 + (b_{21}^j)^2 \rfloor}{\sum_{k=1}^p \sum_{j=1}^q \sum_{lm=1}^2 \left[(a_{lm}^k)^2 + (b_{lm}^j)^2 \right]}$$

Achtung! Floor im
Zähler ist KLAMMER!

Nenner dient der Normierung:
Wert zwischen 0 und 1

Mehrdimensionale dynamische Spektralanalyse auf Basis eines ARMA(p,q) - Modells



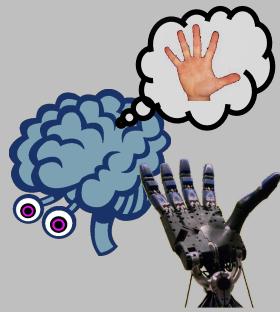
Adaptive Schätzung der Transferfunktion:

$$H_n(\lambda) = A_n^{-1}(\lambda) * B_n(\lambda)$$

$$A_n(\lambda) = I + \sum_{k=1}^p \hat{A}_n^k(n) e^{-ik\lambda}$$

$$B_n(\lambda) = I - \sum_{j=1}^q \hat{B}_n^j(n) e^{-ij\lambda}$$

Kovarianzmatrix



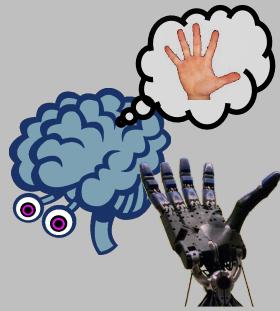
Schätzung der Spektraldichtematrix benötigt Schätzung der **Kovarianzmatrix** des Vorhersagefehlers (Rauschen):

$$s_{ij} = 0$$

$$s_{ij}(n) = s_{ij}(n-1) - c_s \left[s_{ij}(n-1) - (e_i(n) - M_{e_i}^{C_s}(n)) * (e_j(n) - M_{e_i}^{C_s}(n)) \right],$$

$$n=1,2,\dots \text{ mit } i,j=1,\dots,d$$

Spektraldichtematrix



Schätzung der Spektraldichtematrix fortlaufend in der Zeit:

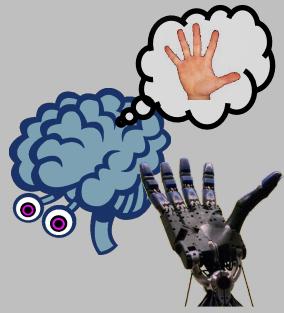
$$f_n(\lambda) = H_n(\lambda) * S_n * H_n^T(\lambda)$$

Zeit – Leistungs - Darstellung der i-ten Komponente:

$$f_{ii}(\lambda_k, n) = f_{ii,n}(\lambda_k), \quad \lambda_k = \frac{2\pi k}{N},$$

$$k=0, \dots, \frac{N}{2}-1, \quad n=0, 1, \dots$$

Spektraldichtematrix



Schätzung der Spektraldichtematrix fortlaufend in der Zeit:

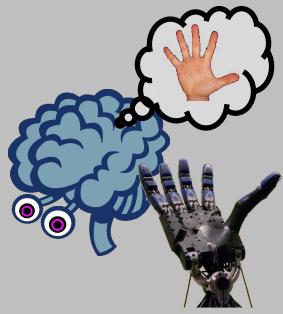
$$f_n(\lambda) = H_n(\lambda) * S_n * H_n^{iT}(\lambda)$$

Zeit – Leistungs - Darstellung der i-ten Komponente:

$$f_{ii}(\lambda_k, n) = f_{ii,n}(\lambda_k), \quad \lambda_k = \frac{2\pi k}{N},$$

$$k=0, \dots, \frac{N}{2}-1, \quad n=0, 1, \dots$$

Kohärenz und Phase-Locking



Zeit – Kohärenz – Darstellung zweier Komponenten i, j :

$$\rho_{i,j}^2(\lambda_k, n) = \frac{|f_{ij,n}(\lambda_k)|^2}{f_{ii,n}(\lambda_k) * f_{jj,n}(\lambda_k)}$$

Zeit – Kreuzphasen – Darstellung zweier Komponenten i, j :

$$\phi_{i,j}(\lambda_k, n) = \arg \left(\Re \{ f_{ij,n}(\lambda_k) \} + i \Im \{ f_{ij,n}(\lambda_k) \} \right)$$