

Die Group Method of Data Handling – eine Verwandte der neuronalen Netze ?

Prof. Dr.-Ing. Tatjana Lange
Fachhochschule Merseburg
FB Elektrotechnik

Beispiele für Anwendungen der GMDH (1):

Einsatzfeld	Ziel	Ausgangsgrößen	Eingangsgrößen	Anwendung
Wirtschaft/ Finanz- wirtschaft	Prognose von monatlichen Zinssätzen	<ul style="list-style-type: none"> • Zinssätze der US-Bundesfonds • Gewinne der 5-Jahres-US-Schatzanweisungen 	zeitlich verzögerte monatliche Änderungen in den <ul style="list-style-type: none"> • Zinssätzen der US-Bundesfonds • Gewinnen der 5-Jahres-US-Schatzanweisungen • Geldumläufen • Indizes der Dollarkurse • Umsätzen • Verbraucherpreisen • industriellen und kommerziellen Anleihen 	Weltbank Washington D.C.
Wirtschaft / Ökologie	Prognose des jährl. Fangertrages in verschiedenen Fangzonen der US-Garnelenfischerei	<ul style="list-style-type: none"> • Jahresfangertrag in 21 statistischen Fangbereichen (Prognose wird zur Jahresmitte erstellt)	<ul style="list-style-type: none"> • Mittel der Nord- und Ostkomponenten des Ekman-Transports (der windgetriebenen Meeresströmung) im ersten und zweiten Quartal • Gesamtwassermenge der Flüsse Mississippi, Sabine, Trinity und Neches im ersten und zweiten Quartal • Minima, Maxima und Mediane der Oberflächenwassertemperatur und des Salzgehalts bei Nord-Jetty 	Nationaler Marine-Fischereidienst der USA

Beispiele für Anwendungen der GMDH (2):

Einsatzfeld	Ziel	Ausgangsgrößen	Eingangsgrößen	Anwendung
Wirtschaft	Industrielle Umsatzprognose	<ul style="list-style-type: none"> Vorauswahl von Einflußfaktoren 	<ul style="list-style-type: none"> 22 mögliche Einflußfaktoren auf den Umsatz ↗ 	Matsushita Electric Works, Osaka, Japan
		<ul style="list-style-type: none"> Umsatz, bezogen auf bestimmte Waren 	Fünf vorselektierte Einflußfaktoren: <ul style="list-style-type: none"> Jahresumsatzplan der Betriebseinheit Umsatzplan für die nächste Periode Erfüllungsrate des betrieblichen Jahresumsatzplanes Warenpreis Warenmangel 	
Ökologie	Kurzzeitprognose der Luftverschmutzung	<ul style="list-style-type: none"> Schwefeldioxyd-Belastung 	<ul style="list-style-type: none"> Schwefeldioxyd-Werte, gemessen während der letzten N Tage ($N = 5, 10, 31$) 	Prognose der Luftverschmutzung in Tokushima, Japan
Ökologie / Meteorologie	Hochwasserprognose	<ul style="list-style-type: none"> Niederschlagsmenge in bestimmten sensiblen Gebieten 	<ul style="list-style-type: none"> Länge und Breite des Taifunzentrums atmosphärischer Druck im Taifunzentrum Ausbreitungsgeschwindigkeit des Taifunzentrums Radius des Taifun-Gebietes Abstand vom Taifunzentrum u. s. w. (insgesamt 11 Eingänge) 	Hochwasserkontrolle im Becken des Flusses Kino, Japan

Beispiele für Anwendungen der GMDH (3):

Clusterisierung und Mustererkennung

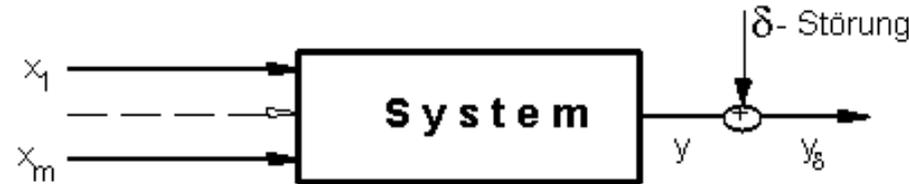
Einsatzfeld	Ziel	Ausgangsgrößen	Eingangsgrößen	Anwendung
Medizin	Diagnose in der Onkologie	<ul style="list-style-type: none"> Diagnose-Trennregeln 	<ul style="list-style-type: none"> nukleotrone Aktivität des Blutserums von krebserkrankten und gesunden Personen bei verschiedenen Verdünnungsgraden 	Krebsdiagnose im Kiewer Onkologischen Zentrum
Metallurgie	Modellierung von Hochofenprozessen	<ul style="list-style-type: none"> Heißwindmenge für jede Windform pro Zeiteinheit 	<ul style="list-style-type: none"> Gesamt-Heißwindmenge vor der Heißwindringleitung Druckdifferenz an den Windformen 	Donezker metallurgisches Werk

Prognose mit neuronalen GMDH-Netzen

Einsatzfeld	Ziel	Ausgangsgrößen	Eingangsgrößen	Anwendung
Finanzwirtschaft	kurzfristige Prognose der Wertpapier-Aktivitäten	<ul style="list-style-type: none"> Aktivitätsindex 	<ul style="list-style-type: none"> Minimum des Aktivitätsindex im zurückliegenden Zeitraum von 70 Tagen Maximum des Aktivitätsindex im zurückliegenden Zeitraum von 70 Tagen Wert des Aktivitätsindex bei Börsenschluß Summarischer Umfang der täglichen Operationen 	New Yorker Wertpapier-Börse

Gegeben sind:

- die gemessenen Eingangsgrößen^l x_j und die gestörte Ausgangsgröße y_δ (Datentabelle):



- Datentabelle:

	$x_0 \equiv 1$	x_1	x_2	...	x_j	...	x_m	y_δ
1					
2					
⋮				⋮		⋮		
i				⋮		⋮		
⋮				⋮		⋮		
l					

- der Modellansatz: (Kolmogorov-Gabor-Polynom):

$$y = a_0 + a_1 x_1 + \dots + a_j x_1^j + \dots + a_n x_1^i x_2^j \dots x_m^u + \dots + a_M x_m^K$$

m - Anzahl der Eingangsgrößen

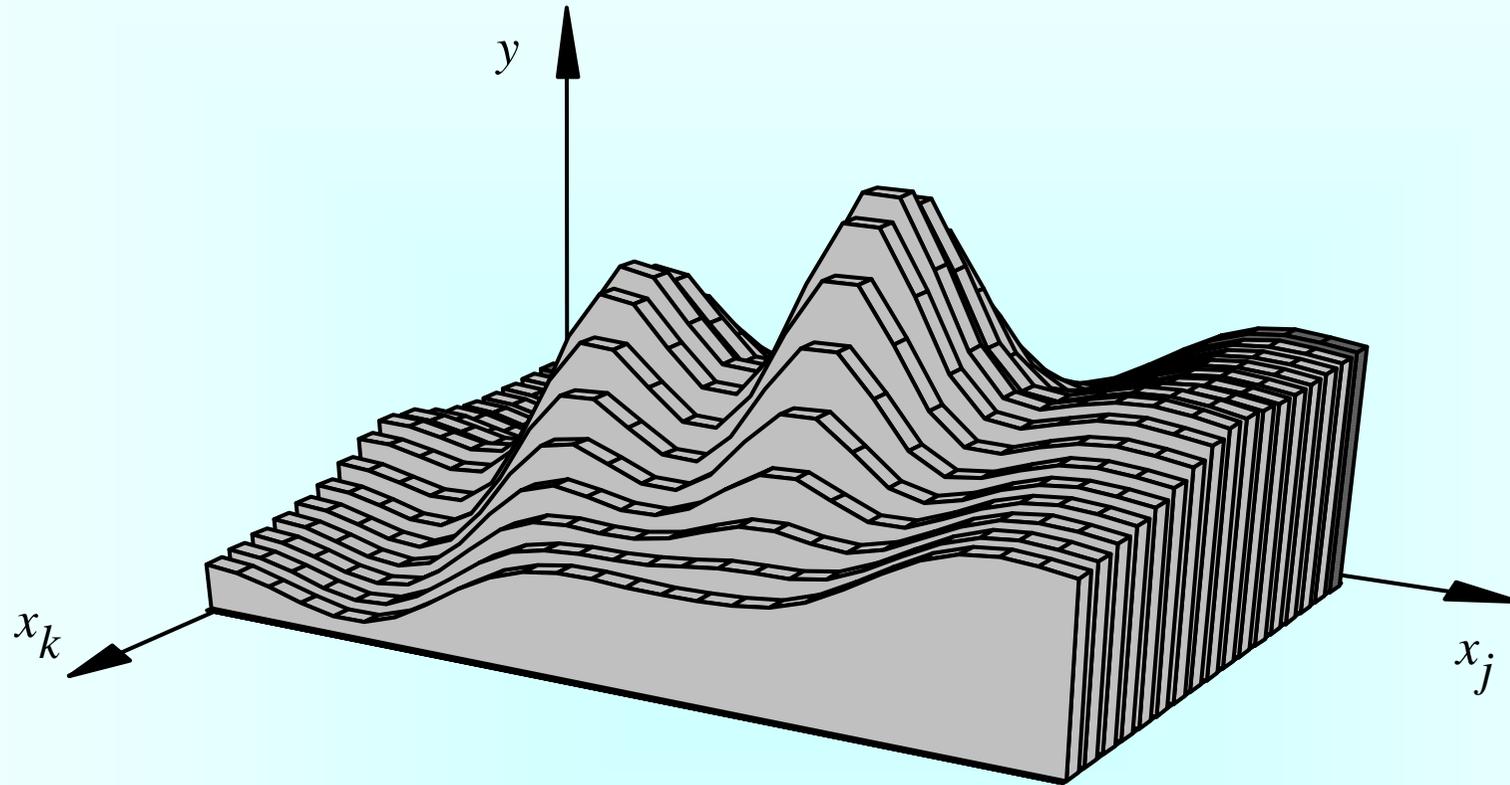
$(M+1)$ - Anzahl der zu schätzenden Parameter

K - maximaler Grad des Polynoms

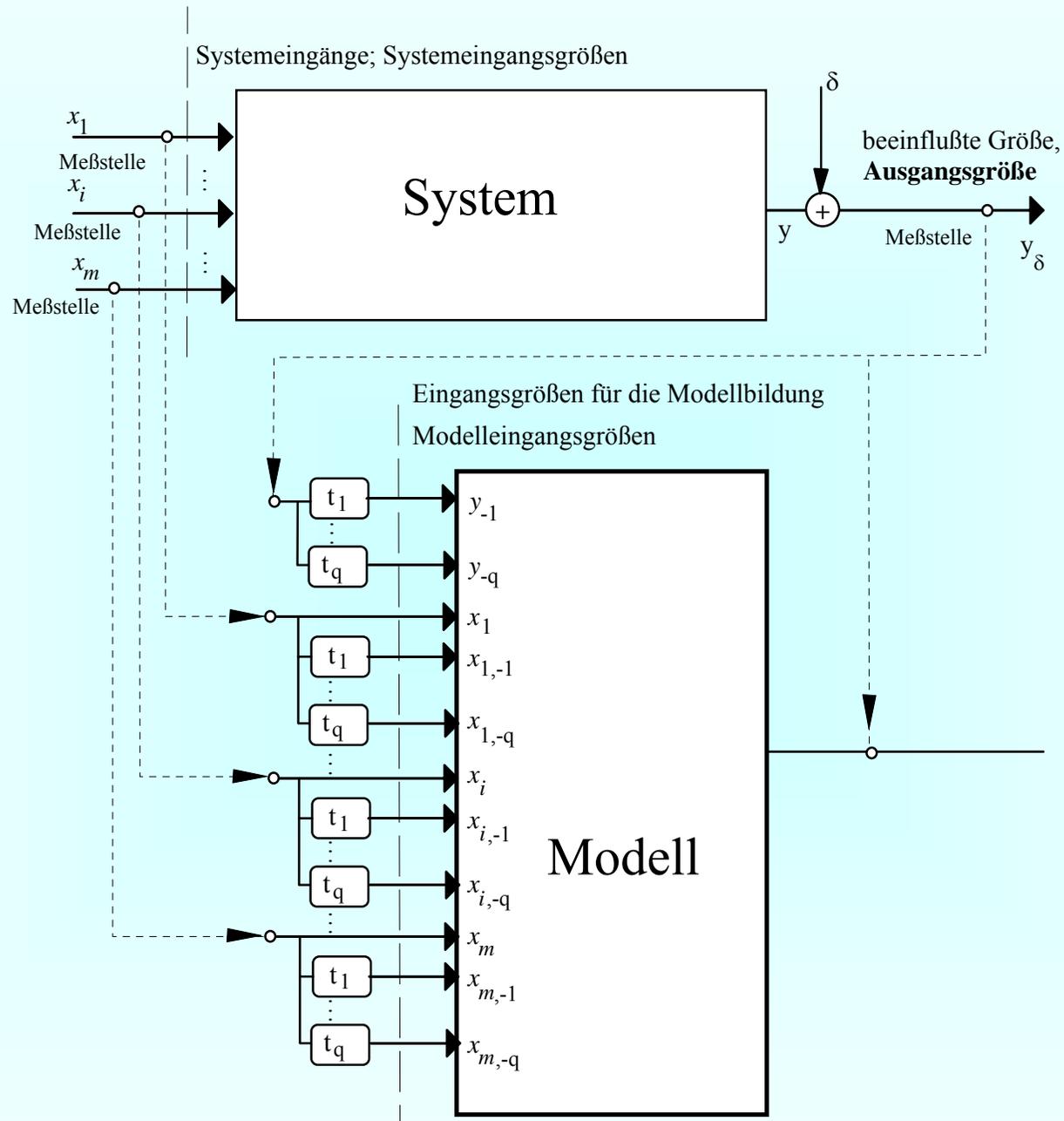
Gesucht sind:

- \vec{a}^* optimale Werte des Parametervektors $\vec{a}^* = (a_0, a_1, \dots, a_M)$

- M^*, K^* optimale Struktur:

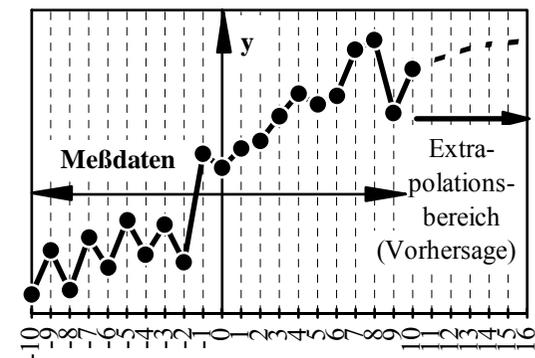
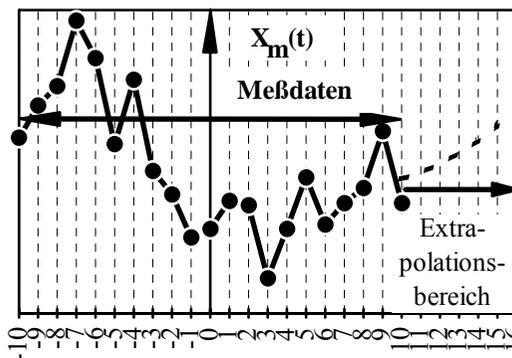
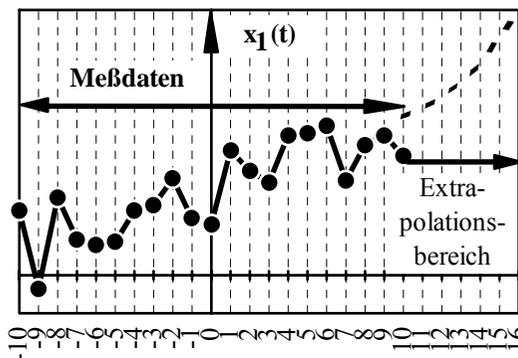


Regressionsfläche



Meßdaten (Beispiel):

$x_{1,k}$	$x_{1,k-1}$	$x_{1,k-2}$...	$x_{1,k-q}$	$x_{m,k}$	$x_{m,k-1}$	$x_{m,k-2}$...	$x_{m,k-q}$	Nr	y_k	y_{k-1}	y_{k-2}	...	y_{k-q}
$x_1(0)$	$x_1(-1)$	$x_1(-2)$			$x_m(0)$	$x_m(-1)$	$x_m(-2)$			0	$y(0)$	$y(-1)$	$y(-2)$		
$x_1(1)$	$x_1(0)$	$x_1(-1)$			$x_m(1)$	$x_m(0)$	$x_m(-1)$			1	$y(1)$	$y(0)$	$y(-1)$		
$x_1(2)$	$x_1(1)$	$x_1(0)$			$x_m(2)$	$x_m(1)$	$x_m(0)$			2	$y(2)$	$y(1)$	$y(0)$		
.
.
.
$x_1(10)$	$x_1(9)$	$x_1(8)$			$x_m(10)$	$x_m(9)$	$x_m(8)$			10	$y(10)$	$y(9)$	$y(8)$		



Alle beobachtbaren Größen, die auf die interessierende Größe Einfluß haben können, sollen nachfolgend **Systemeingangsgrößen** genannt werden.

Die Größe, die von ihnen beeinflusst wird und die zu steuern oder zu prognostizieren ist, soll als **Ausgangsgröße** bezeichnet werden. Auch diese Größe sei beobachtbar.

Dem betrachteten System steht nun das Modell gegenüber, das **Modelleingangsgrößen** mit der **Ausgangsgröße** verbindet.

Im einfachsten Fall sind die Systemeingangsgrößen und die Modelleingangsgrößen identisch.

Um die Anwendbarkeit der Modellierungsmethode zu erweitern, ist es jedoch sinnvoll, zunächst die Modelleingangsgrößen von den Systemeingangsgrößen zu unterscheiden.

Verallgemeinert können dann diese Modelleingangsgrößen aus am System aufgenommenen Meßreihen der Systemeingangsgrößen und der Ausgangsgröße unterschiedlich gebildet werden (s. unten). Einen möglichen Ansatz dafür zeigt Abb. 1.4.a.

Im allgemeinen wird auch der Begriff **Einflußgrößen** verwendet. Darunter sollen alle auf das System einwirkende Größen verstanden werden, also auch die, die nicht gemessen werden können und die, die nicht bekannt oder nicht beobachtbar sind.

Es seien nun drei praktische Anwendungsgruppen betrachtet, bei denen die am System gemessenen Eingangsgrößen (Systemeingangsgrößen) auf unterschiedliche Art und Weise auf die Eingangsgrößen für die Modellbildung (Modelleingangsgrößen) abgebildet werden.

Erste Gruppe:

Die Modelleingangsgrößen spiegeln nur die unmittelbare Wirkung der Systemeingangsgrößen auf die interessierende, zu prognostizierende Größe (Ausgangsgröße, beeinflusste Größe) wider. Etwaige verzögerte oder sich im System träge auswirkende Systemeingangsgrößen werden nicht berücksichtigt. In diesem Fall sind Modelleingangsgrößen und Systemeingangsgrößen identisch.

Die Abhängigkeit, die durch die Regression dargestellt wird, ist in diesem Fall eine **statische Charakteristik** (bzw. statische Kennlinie), die keine Zeitverzögerungen berücksichtigt.

Statische Kennlinien werden für die Beschreibung von Systemen mit kleinen Speichern (Energie-, Stoffspeicher u.ä.) verwendet, deren Kapazität bei der Berechnung vernachlässigt werden kann (statische Systeme, d.h. Systeme ohne Trägheit).

Meßdaten, die für die Rekonstruktion statischer Kennlinien mit Hilfe der Regression notwendig sind, sind in diesem Fall Meßreihen der Systemeingangsgrößen und der beeinflussten Größe (Systemausgangsgröße).

Bei der Rekonstruktion der statischen Kennlinie von dynamischen Systemen werden Meßreihen der gleichen Größen wie im Fall statischer Systeme benötigt, die jedoch erst nach Abklingen der Übergangsprozesse aufgenommen werden dürfen.

Zweite Gruppe:

Die Modelleingangsgrößen spiegeln sowohl die unmittelbare Wirkung als auch die verzögerte Wirkung von Systemeingangsgrößen auf die zu prognostizierende Größe wider sowie die Wirkung der zu prognostizierenden Größe (Ausgangsgröße) auf sich selbst. Die Modelleingangsgrößen spiegeln also auch die Trägheit im System wider, die etwa durch große Speicher im System verursacht wird.

Die Abhängigkeit
$$y_k = f(y_{k-1}, y_{k-2}, \dots, y_{k-q}, x_{1,k}, x_{1,k-1}, \dots, x_{1,k-q}, x_{2,k-1}, \dots, x_{2,k-q}, \dots, x_{m,k-1}, \dots, x_{m,k-q})$$

die durch die Regression dargestellt wird, ist in diesem Fall eine dynamische **Charakteristik** (dynamische Kennlinie), d.h. die Lösung der Differential- bzw. Differenzgleichungen, die das System beschreiben, für das entsprechende Eingangssignal (Meßreihe der Werte der Systemeingangsgrößen).

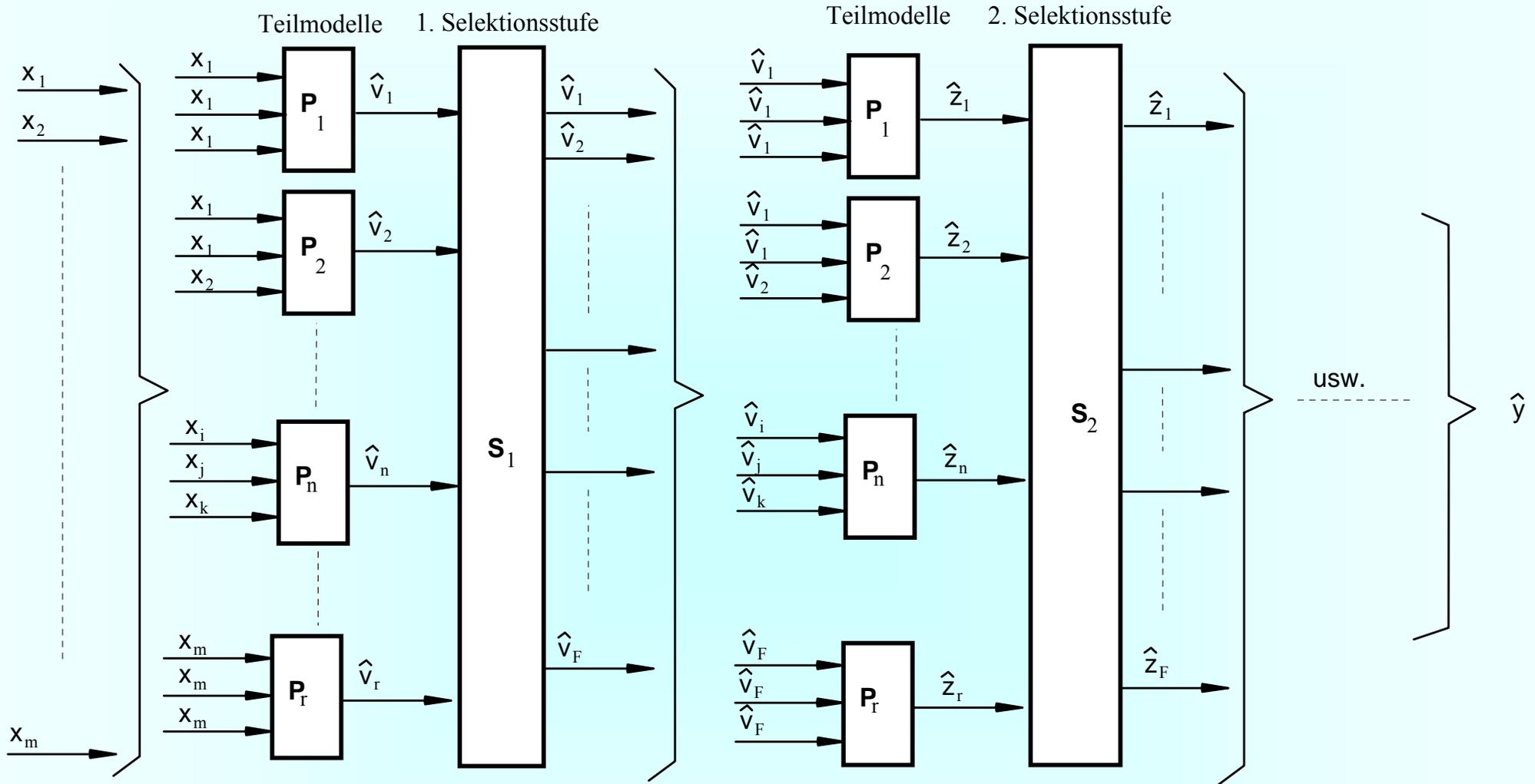
Dritte Gruppe.

Die Modelleingangsgrößen spiegeln nur die Verzögerung des zu prognostizierenden Prozesses (der Ausgangsgröße) wider. Die Abhängigkeit, die durch die Regression beschrieben wird, ist in diesem Fall eine **Autoregression**.

In Abhängigkeit davon, zu welcher Gruppe die Anwendungsaufgabe gehört, erfolgt die Bildung der Eingangsmatrix für die Modellierung aus den vorhandenen Meßdaten (s. Tab. 1.4.a und Abb. 1.4.b).

Statische und dynamische Ein-/Ausgangsmodelle

		Anwendungsbereich		
	Bestimmung von Modellen statischer Kennlinien als Funktion $y = f(x_1, x_2, \dots, x_m)$	\Rightarrow dynamischer Kennlinien nichtlinearer Systeme \Rightarrow mehrdimensionaler Zeitreihen als Funktion $y_k = f(y_{k-1}, y_{k-2}, \dots, y_{k-q}, x_{1,k}, x_{1,k-1}, \dots, x_{1,k-q}, x_{2,k-1}, \dots, x_{2,k-q}, \dots, x_{m,k-1}, \dots, x_{m,k-q})$	Bestimmung von Modellen für den Autoregressionsprozeß als Funktion $y_k = f(y_{k-1}, y_{k-2}, \dots, y_{k-q})$	
Anzahl der Terme des vollständigen K.-Gabor-Polynoms	$v = \frac{(1+m+K)!}{K!(1+m)!} - 1$	$v = \frac{(1+m+q+K)!}{K!(1+m+q)!} - 1$		$v = \frac{(1+q+K)!}{K!(1+q)!} - 1$
Bildung der Eingangsmatrix \mathbf{X}	$\mathbf{X} = \begin{bmatrix} 1 & x_1(0) & \dots & x_m(0) \\ 1 & x_1(1) & \dots & x_m(1) \\ \dots & \dots & \dots & \dots \\ 1 & x_1(l) & \dots & x_m(l) \end{bmatrix}$	$\mathbf{X} = \begin{bmatrix} y(-1) & y(-2) & \dots & x_1(0) & x_1(-1) & \dots & x_m(-q) \\ y(0) & y(-1) & \dots & x_1(1) & x_1(0) & \dots & x_m(1-q) \\ y(1) & y(0) & \dots & x_1(2) & x_1(1) & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ y(l) & y(l-1) & \dots & x_1(l) & x_1(l-1) & \dots & x_m(l-q) \end{bmatrix}$		$\mathbf{X} = \begin{bmatrix} y(-1) & y(-2) & \dots & y(-q) \\ y(0) & y(-1) & \dots & y(1-q) \\ y(1) & y(0) & \dots & \dots \\ \dots & \dots & \dots & \dots \\ y(l) & y(l-1) & \dots & y(l+1-q) \end{bmatrix}$
Ausgangsvektor	$\vec{y} = \vec{y}_k = \begin{bmatrix} y(0) \\ y(1) \\ \dots \\ y(l) \end{bmatrix}$	$\vec{y}_k = \begin{bmatrix} y(0) \\ y(1) \\ \dots \\ y(l) \end{bmatrix}$		$\vec{y}_k = \begin{bmatrix} y(0) \\ y(1) \\ \dots \\ y(l) \end{bmatrix}$
m - Anzahl der Variablen		q - Anzahl der Verzögerungsschritte		K - Grad des Polynoms



- P_n - n-tes Teilmodell auf der jeweiligen Stufe; Grundform (s.g. Stützfunktion): $v = Ax_i + Bx_j + Cx_k$ bzw. $z = Cv_i + Dv_j + Ev_k$
- \hat{v}_n - auf Basis des Teilmodells P_n geschätzter Ausgang
- S_I - Selektion der F besten Zwischenmodelle aus r Teilmodellen
- \hat{y} - geschätzter Ausgang des endgültigen Modells (ausgang des besten Modells auf der letzten Iterationsstufe)

Beispiel: Modelle mit nichtlinearer Stützfunktion

Die Eingangsmatrix \mathbf{X} enthält 4 Variable

1. Stufe:

$$v_{12} = v_1 = f(x_1, x_2) = a_{11}^{(1)} x_1 + a_{21}^{(1)} x_1 x_2$$

$$v_{13} = v_2 = f(x_1, x_3) = a_{12}^{(1)} x_1 + a_{22}^{(1)} x_1 x_3$$

$$v_{13} = v_3 = f(x_1, x_4) = a_{13}^{(1)} x_1 + a_{23}^{(1)} x_1 x_4$$

$$v_{23} = v_4 = f(x_2, x_3) = a_{14}^{(1)} x_2 + a_{24}^{(1)} x_2 x_3$$

$$v_{24} = v_5 = f(x_2, x_4) = a_{15}^{(1)} x_2 + a_{25}^{(1)} x_2 x_4$$

$$v_{34} = v_6 = f(x_3, x_4) = a_{16}^{(1)} x_3 + a_{26}^{(1)} x_3 x_4$$

2. Stufe:

$$z_{12} = z_1 = f(v_1, v_3) = a_{11}^{(2)} v_1 + a_{21}^{(2)} v_1 v_3$$

$$z_{13} = z_2 = f(v_1, v_6) = a_{12}^{(2)} v_1 + a_{22}^{(2)} v_1 v_6$$

$$z_{13} = z_3 = f(v_3, v_6) = a_{13}^{(2)} v_3 + a_{23}^{(2)} v_3 v_6$$

USW.

⋮

Rückrechnung der Koeffizienten:

Es soll angenommen werden, dass das Abbruchkriterium gezeigt hat, dass sich das optimale Modell auf der 2^{ten} Stufe befindet und das Modell Modell z_3 ist.

$$\begin{aligned}
 z_3 &= a_{13}^{(2)} v_3 + a_{23}^{(2)} v_3 v_6 = \\
 &= a_{13}^{(1)} \underbrace{\left(a_{13}^{(1)} x_1 + a_{23}^{(1)} x_1 x_4 \right)}_{v_3} + a_{23}^{(2)} \underbrace{\left(a_{13}^{(1)} x_1 + a_{23}^{(1)} x_1 x_4 \right)}_{v_3} \underbrace{\left(a_{16}^{(1)} x_3 + a_{26}^{(1)} x_3 x_4 \right)}_{v_6} = \\
 &= a_{13}^{(2)} a_{13}^{(1)} x_1 + a_{13}^{(1)} a_{23}^{(1)} x_1 x_4 + a_{23}^{(2)} a_{13}^{(1)} a_{16}^{(1)} x_1 x_3 + \\
 &\quad + a_{23}^{(2)} a_{23}^{(1)} a_{16}^{(1)} \underline{x_1 x_4 x_3} + a_{23}^{(2)} a_{13}^{(1)} a_{26}^{(1)} \underline{x_1 x_3 x_4} + a_{23}^{(2)} a_{23}^{(1)} a_{26}^{(1)} x_1 x_4 x_3 x_4
 \end{aligned}$$

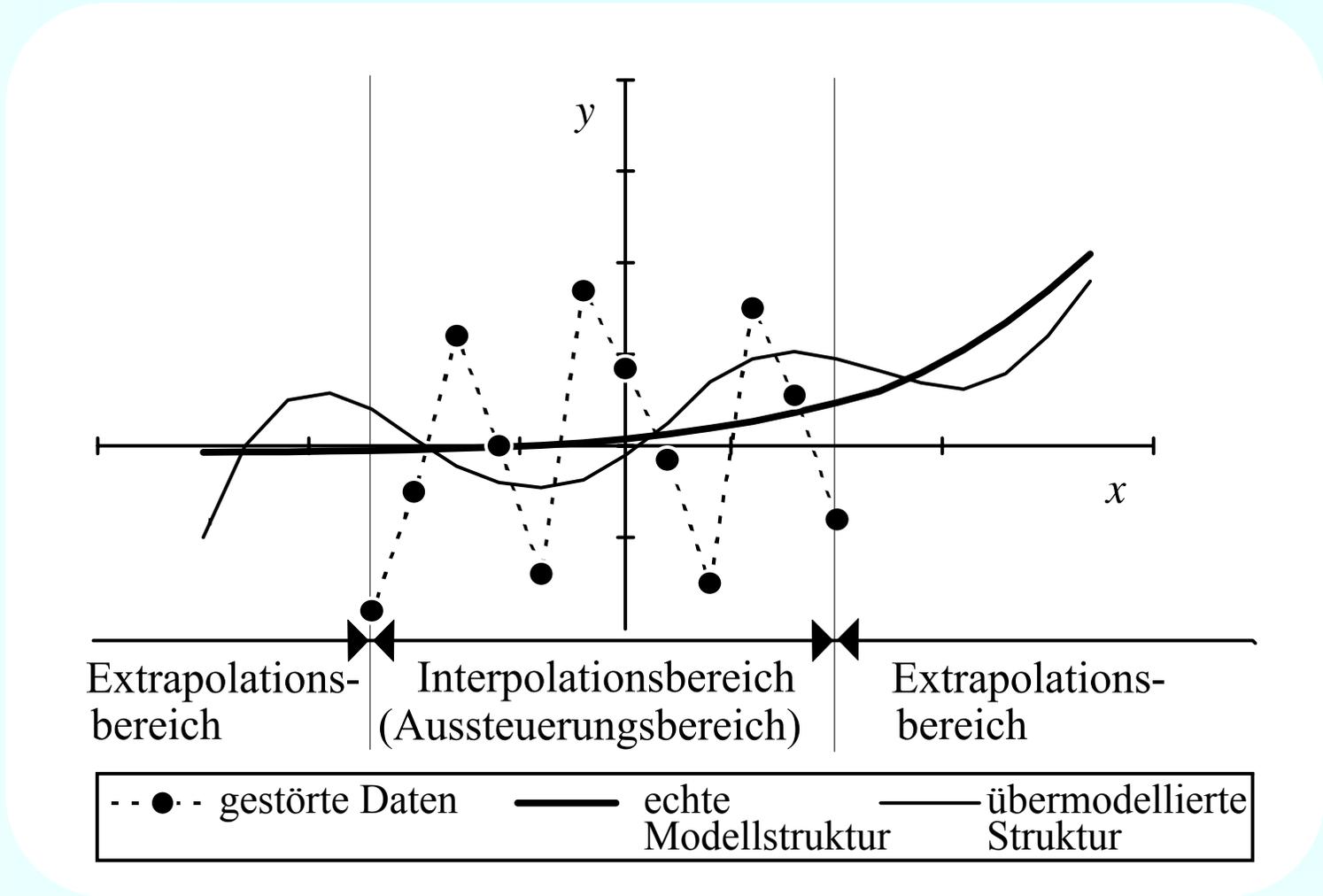
$$z_3 = A_1 x_1 + A_2 x_1 x_4 + A_3 x_1 x_3 + A_4 x_1 x_4 x_3 + A_5 x_1 x_4 x_3 x_4$$

mit

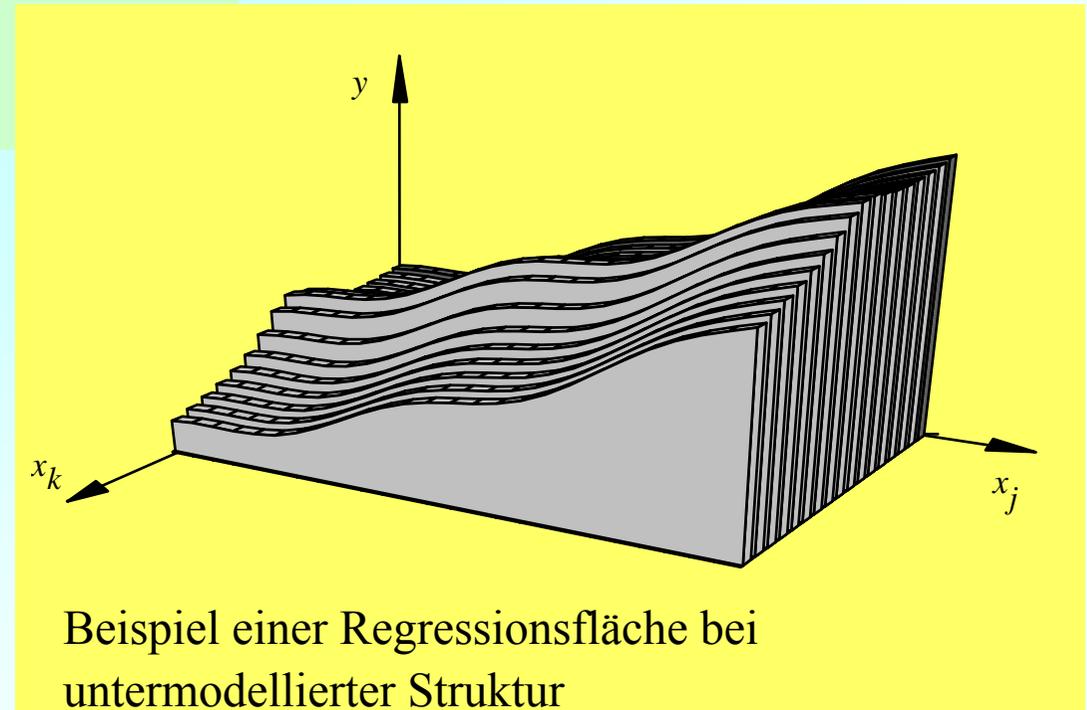
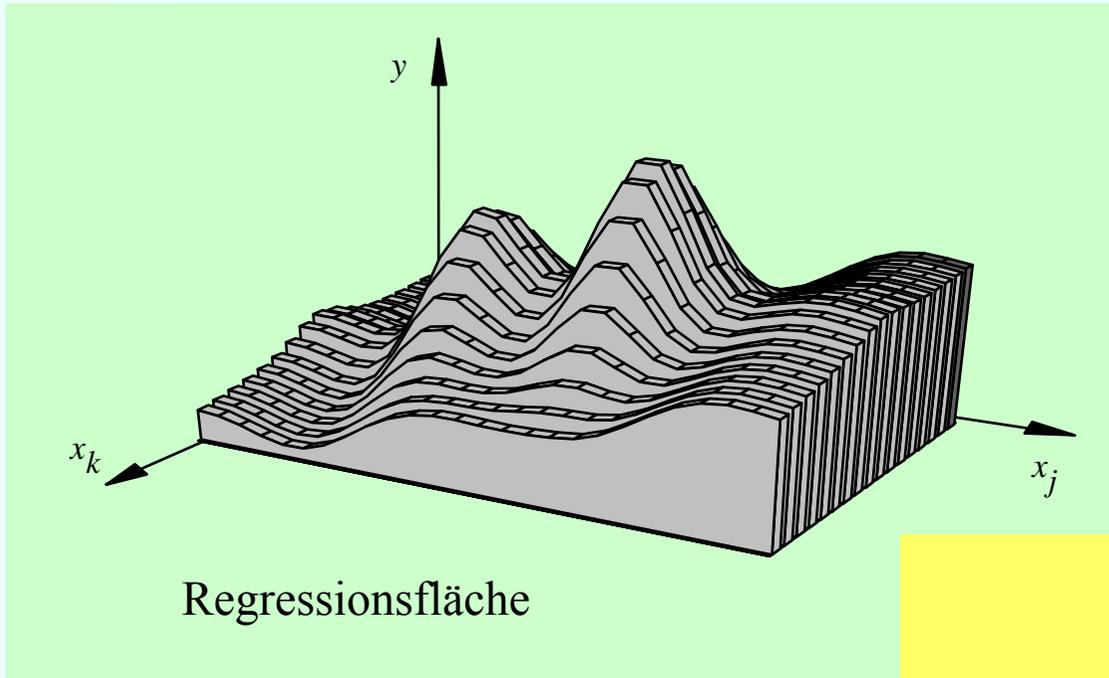
$$\begin{aligned}
 A_1 &= a_{13}^{(2)} a_{13}^{(1)} \\
 A_2 &= a_{13}^{(1)} a_{23}^{(1)} \\
 A_3 &= a_{23}^{(2)} a_{13}^{(1)} a_{16}^{(1)} \\
 A_4 &= a_{23}^{(2)} a_{23}^{(1)} a_{16}^{(1)} + a_{23}^{(2)} a_{13}^{(1)} a_{26}^{(1)} \\
 A_5 &= a_{23}^{(2)} a_{23}^{(1)} a_{26}^{(1)}
 \end{aligned}$$

x_2 wurde durch die Konkurrenz der partiellen Modelle eliminiert !

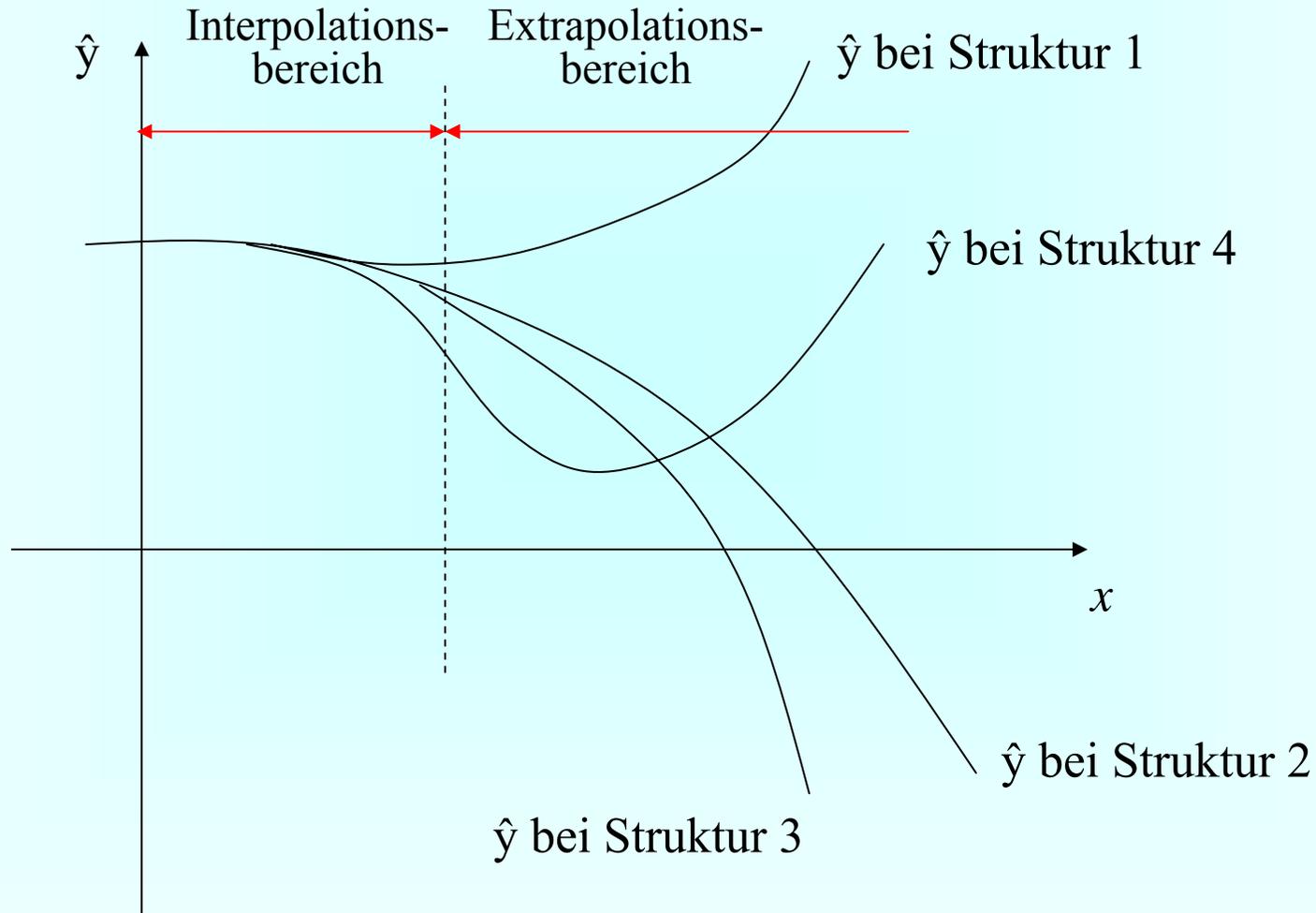
**Problem der Prognosefähigkeit bei der Modellierung mit NN
– vgl. GMDH**



Vergleich einer echten und einer übermodellierten Struktur



Wirkung der Wahl des Kriteriums auf die Prognosefähigkeit:



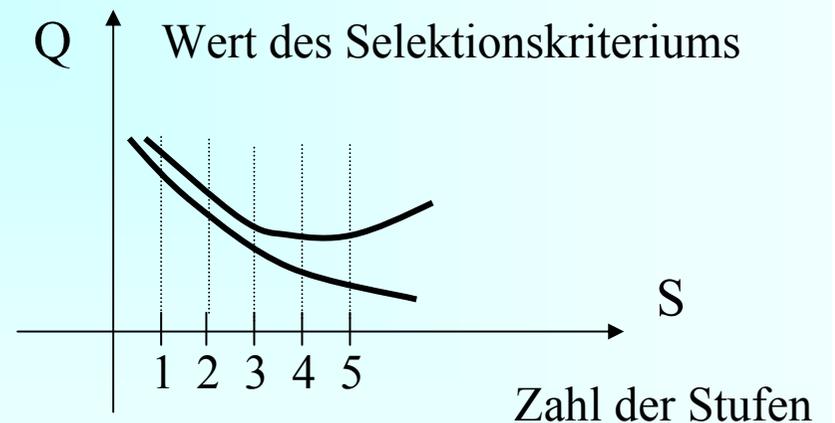
Teilung der Daten in eine Lehrfolge und eine Prüffolge

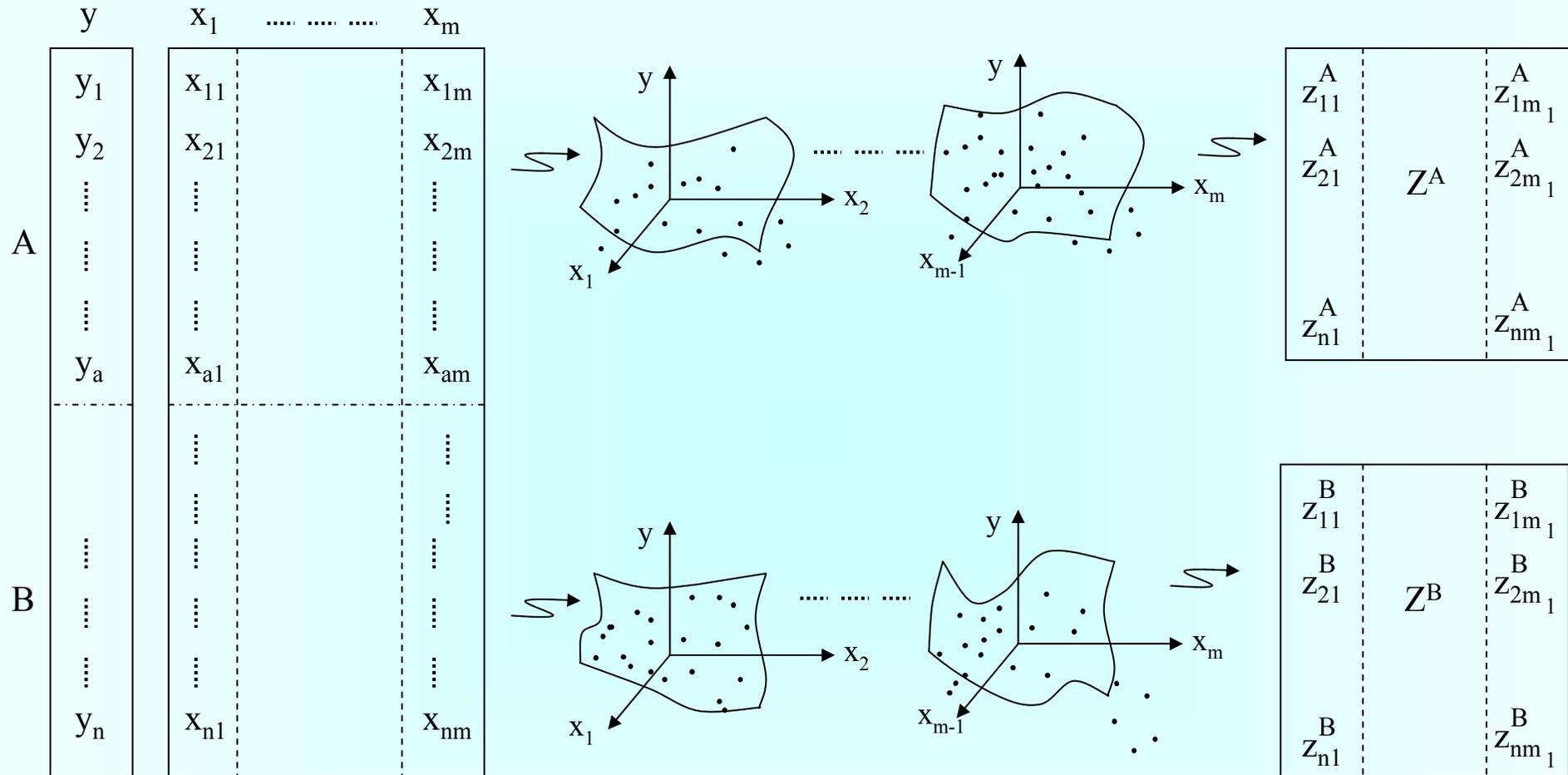
	x_1	x_2	x_m	y
1					
2					
⋮					
⋮					
⋮					
⋮					
l					

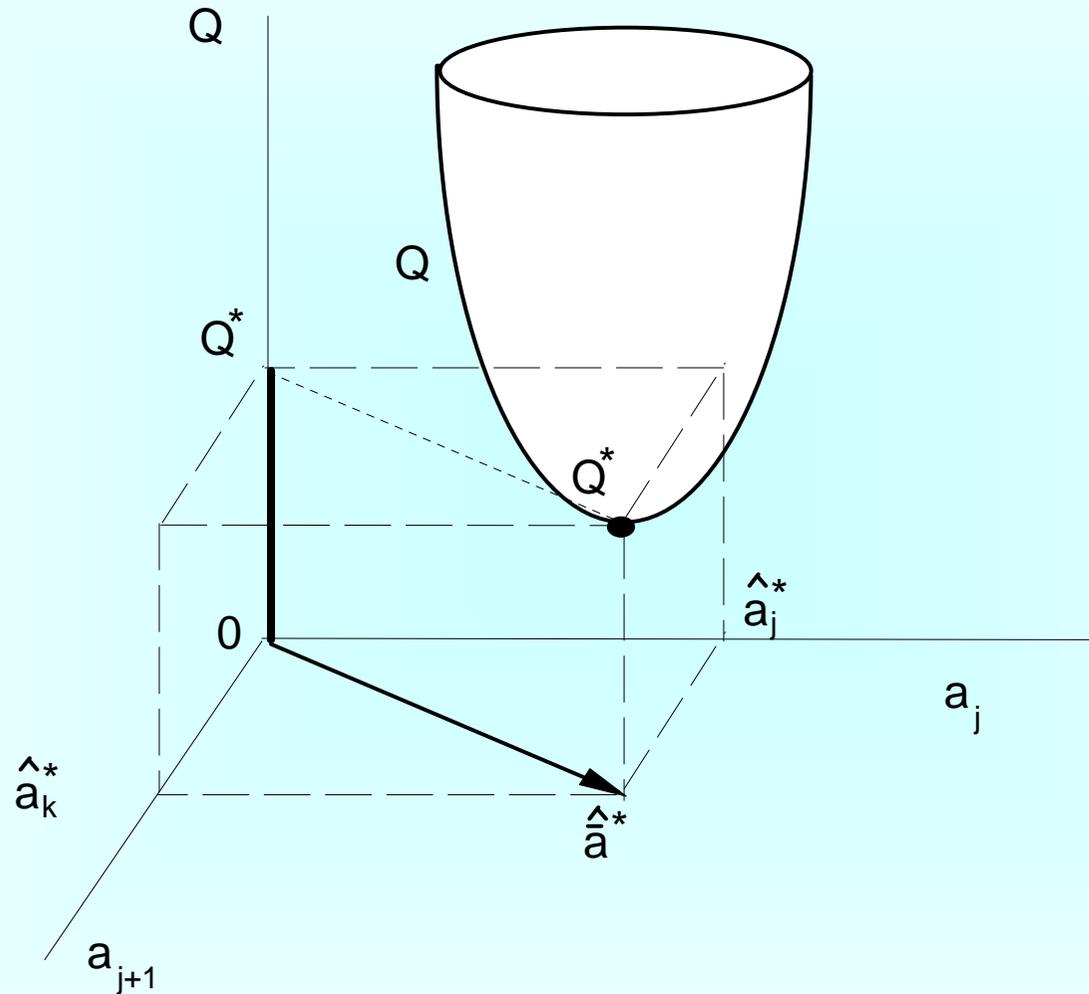
} Lehr- oder Trainingsfolge
} Prüffolge

Die Modelle werden bis zur Erreichung des Minimums des gewählten Selektionskriteriums immer komplizierter.
 Das endgültige Modell könnte beispielsweise folgende Form haben:

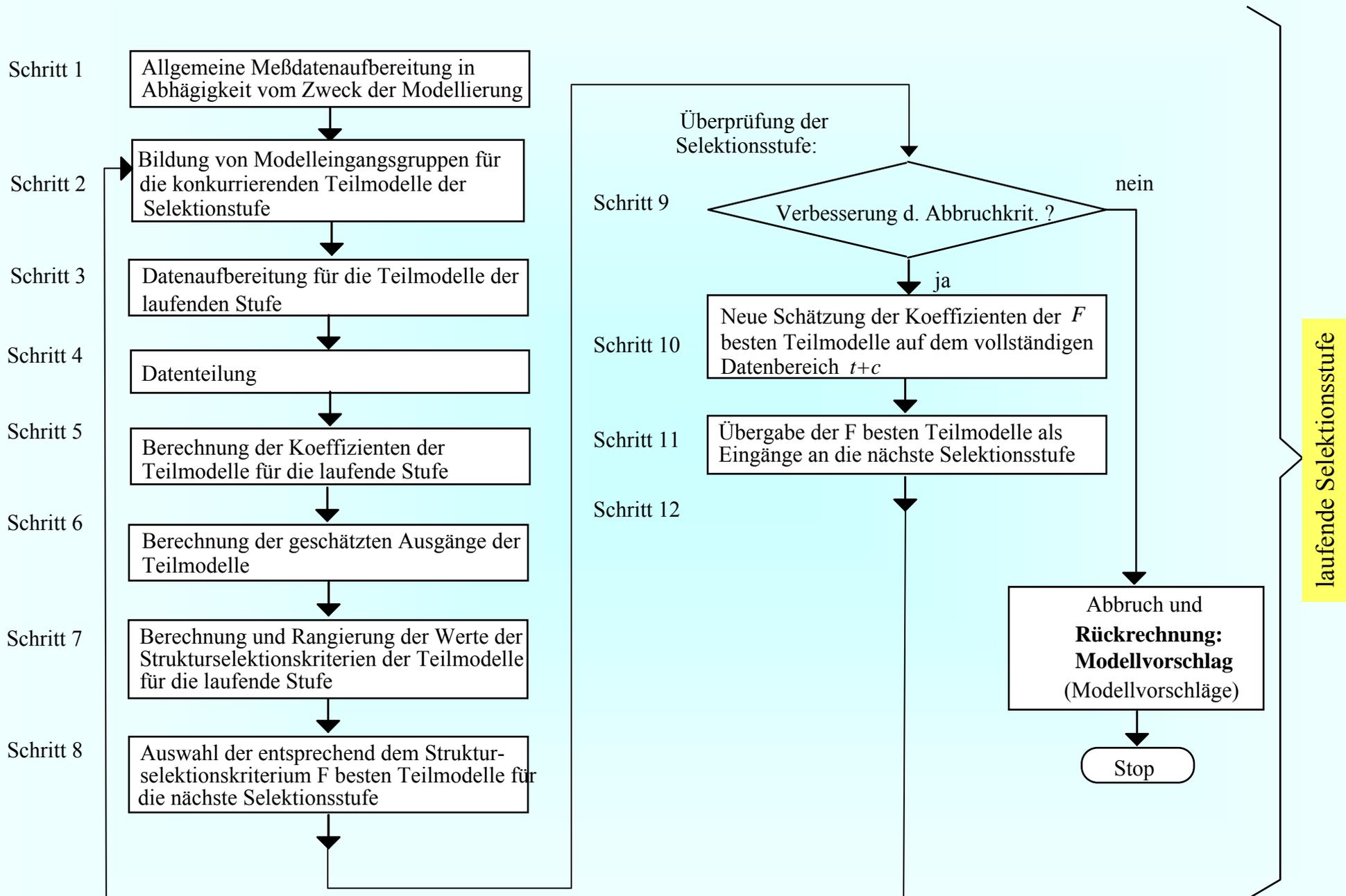
$$y = A_0 + A_1 x_{17} + A_2 x_6 x_9 + \dots + A_i x_1^3 x_2^5 x_3 + \dots$$



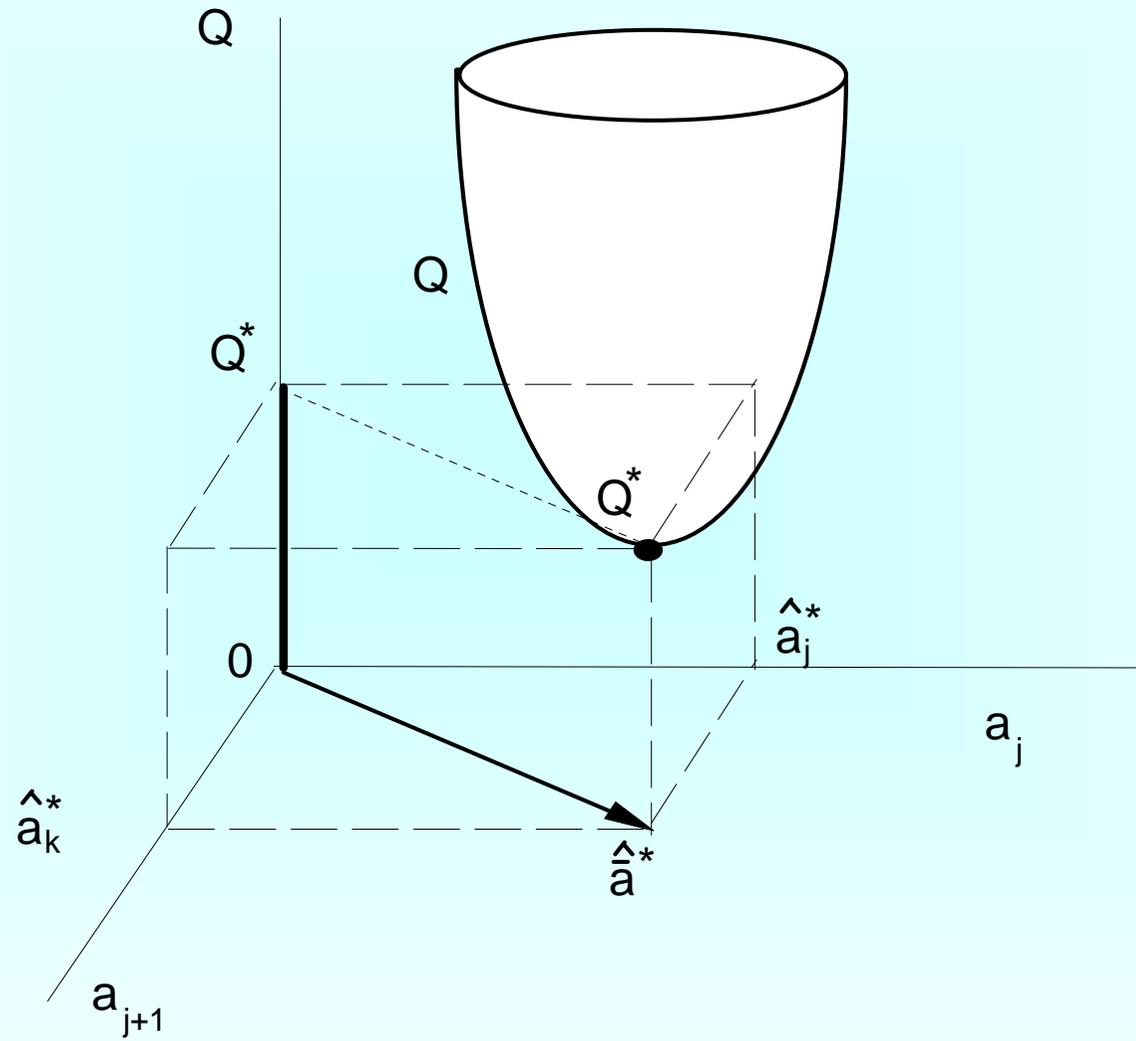


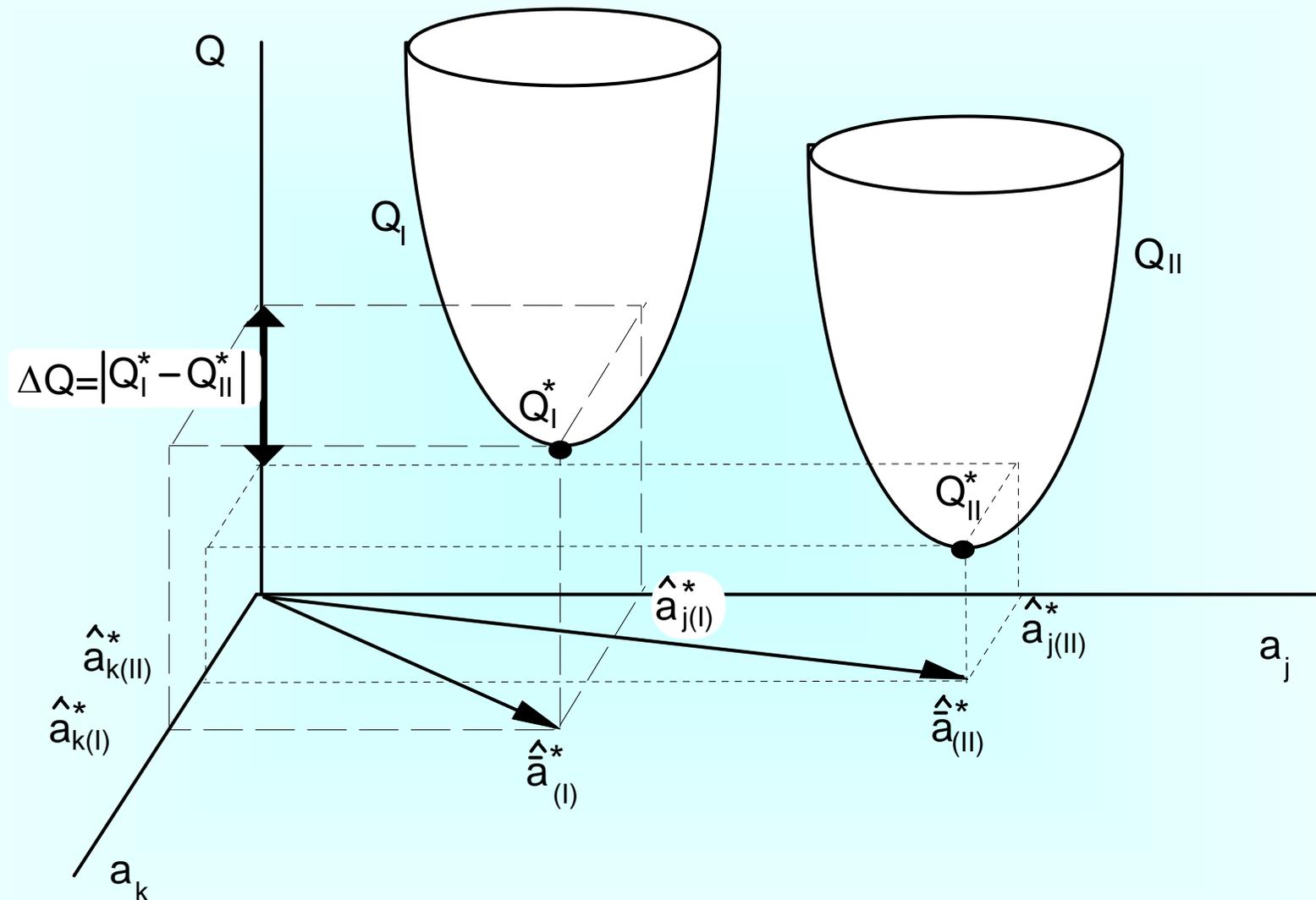


Graphische Darstellung der gesuchten Lösung im Parameterraum

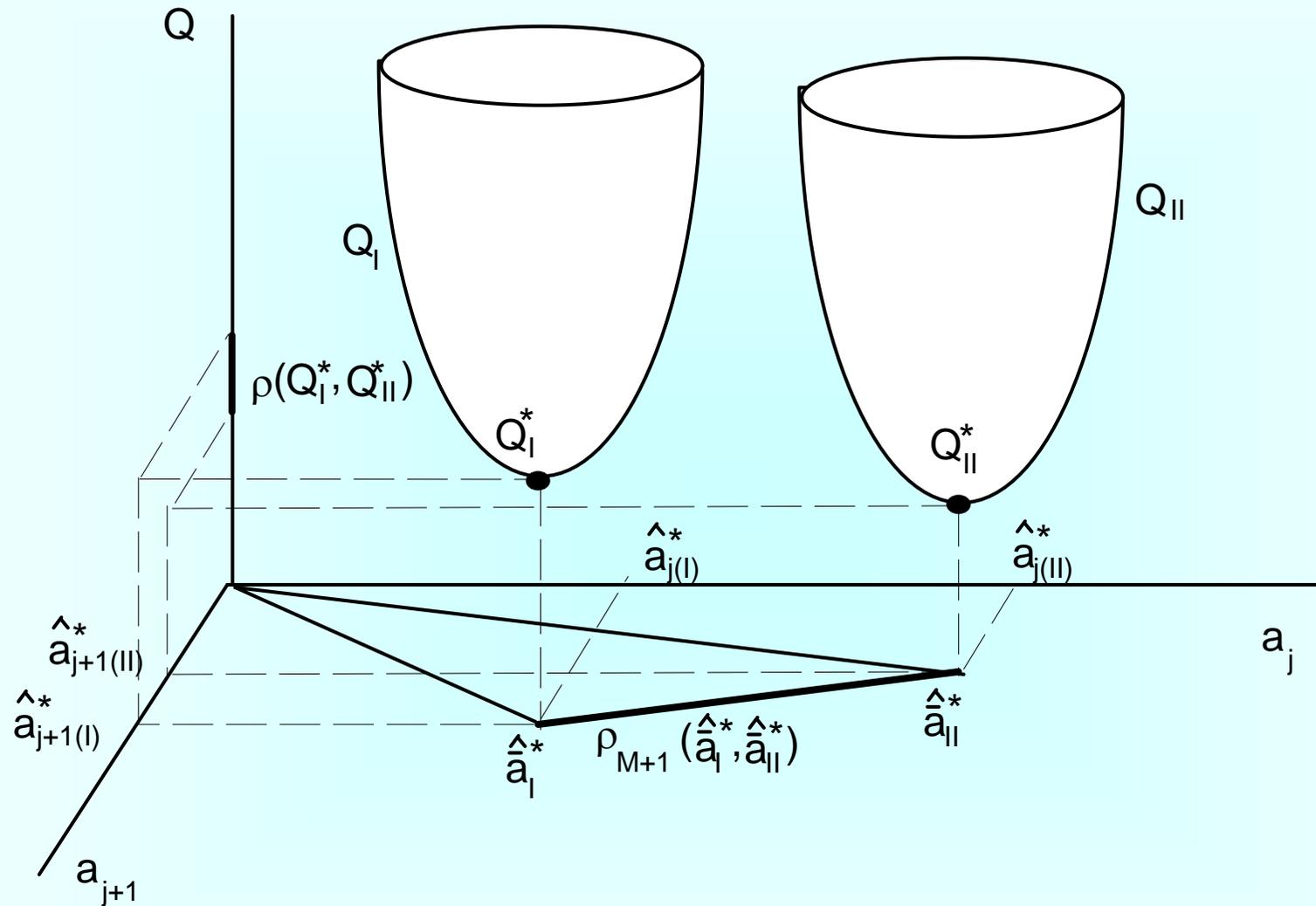


Gemeinsame Probleme der GMDH und der neuronalen Netze bei der Modellierung

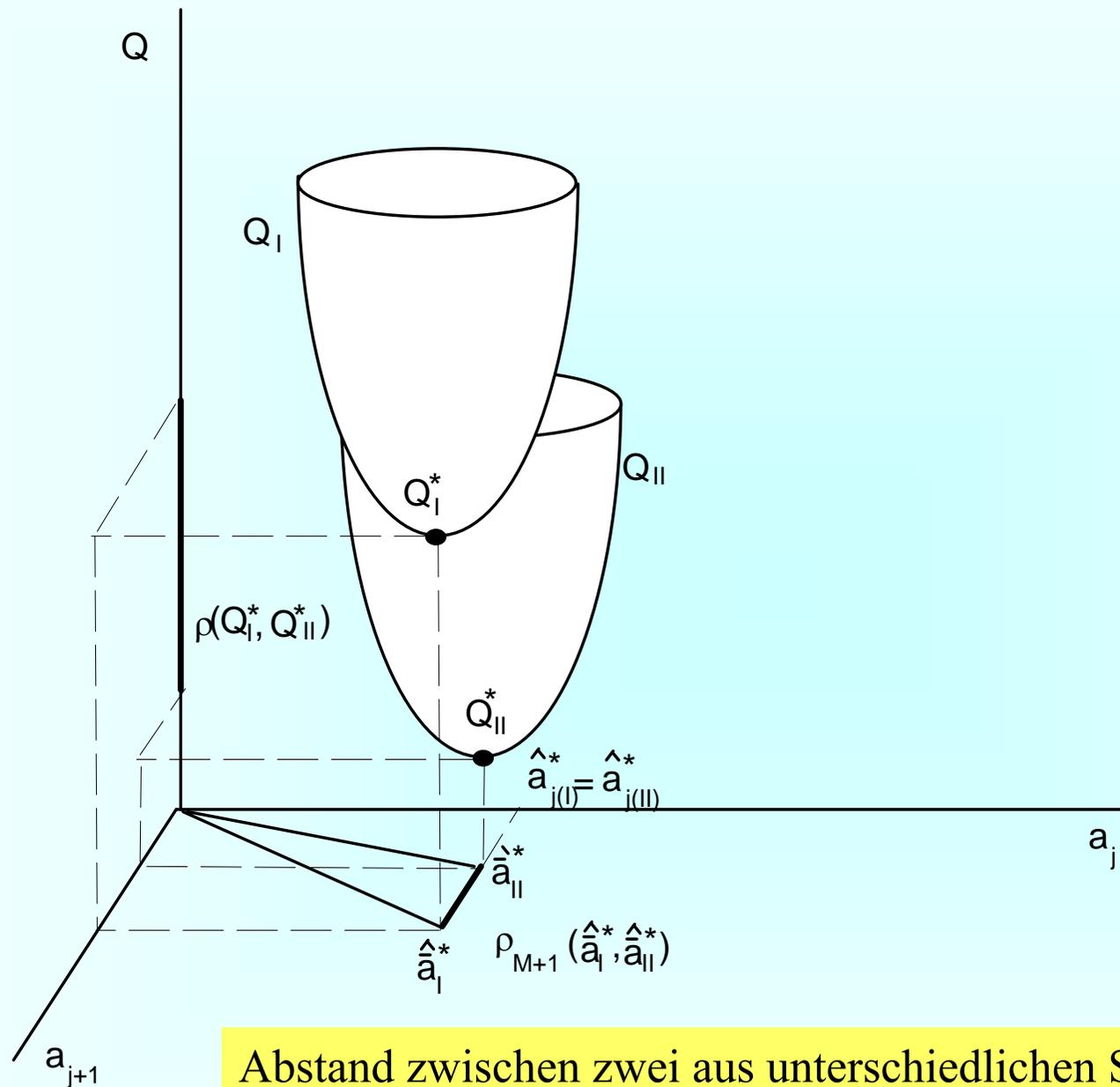




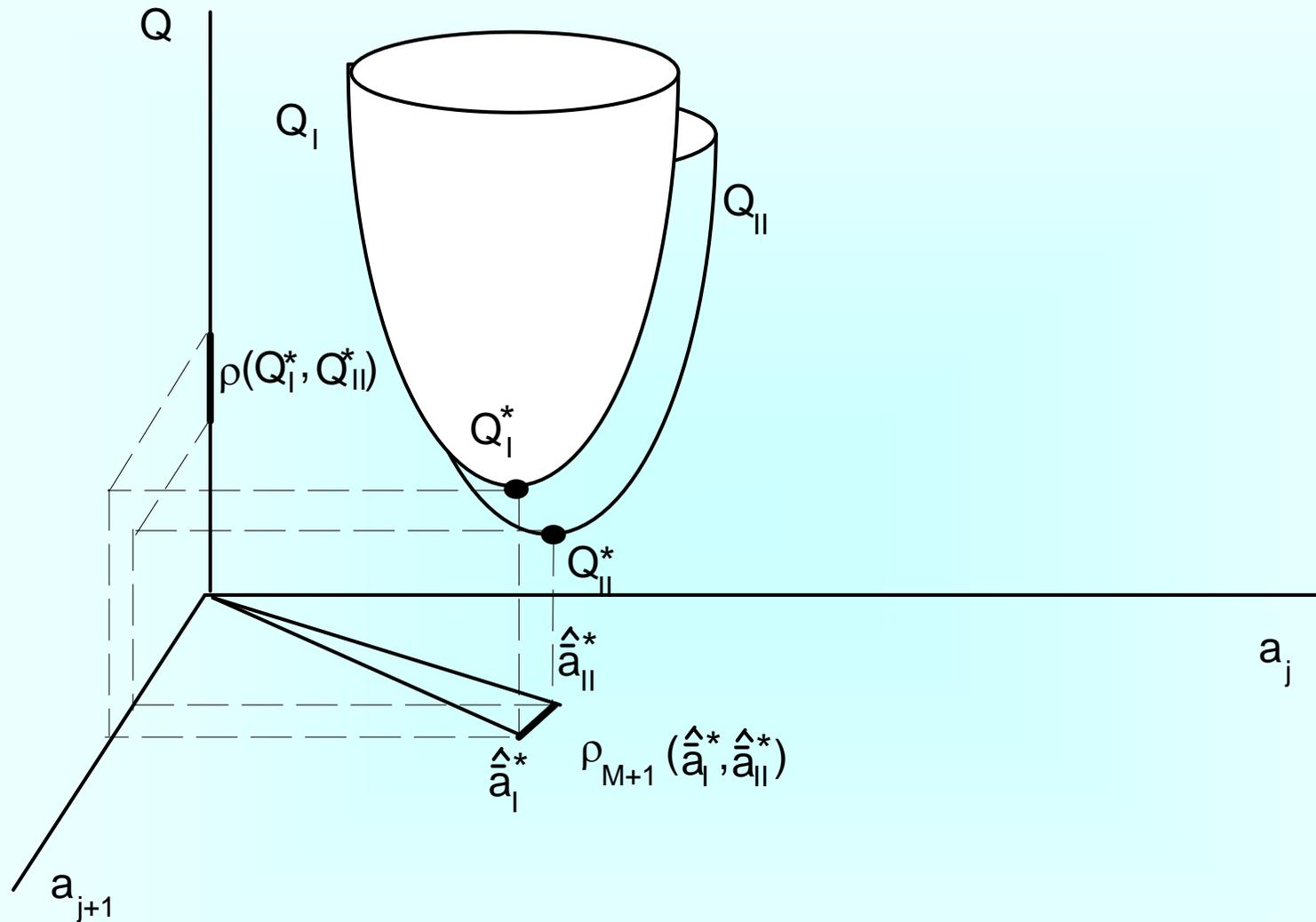
Maß für den Vergleich zweier Strukturen



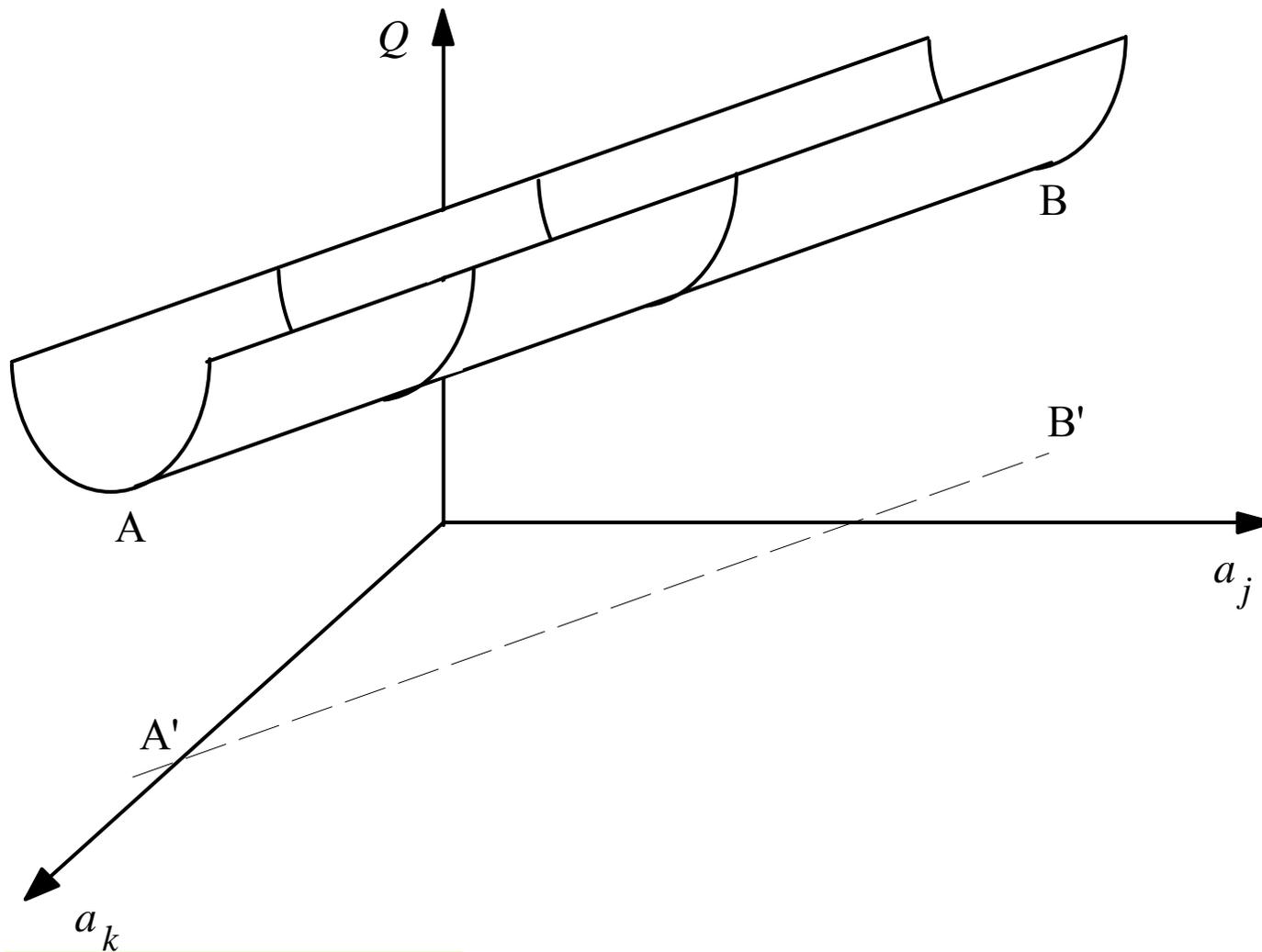
Abstand zwischen zwei aus unterschiedlichen Stichproben gebildeten Modellen
(Nähe der Modelle bezüglich ihres Ausgangs y)



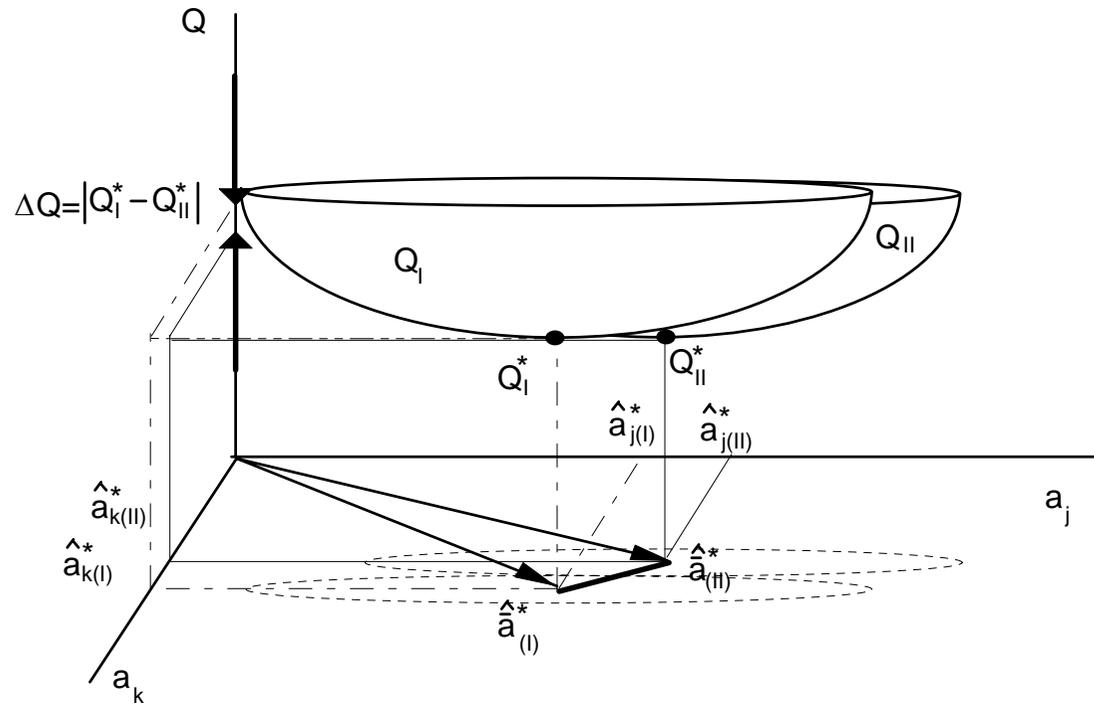
Abstand zwischen zwei aus unterschiedlichen Stichproben gebildeten Modellen (Nähe der Modelle bezüglich der Koeffizienten)



Abstand zwischen zwei aus unterschiedlichen Stichproben gebildeten Modellen
(vollständige Nähe der Modelle)



Instabiles Funktional



Instabilität der Lösung bei extrem flachen Funktionalen

Projektion der Niveaulinien des Gebildes des Parameterkriteriums (Funktionals) im Fall einer instabilen Lösung

