Vortrag zum Thema Regression am 26.9.06 vor Kollegen des Projektes CAS
(Wolfgang Buhmann)

Liebe Kolleginnen und Kollegen,

Regression ist nicht nur ein Gebiet, eine Fachmethode der Mathematik, das in unseren Lehrplänen keine Verbindlichkeit hat, sondern
sie ist zum Einen
ein ideales Beispiel für die sinnvolle Nutzung sog. mathematischer Werkzeuge, da der reine Rechenaufwand ohne diese Hilfsmittel inakzeptabel ist und
zum Anderen
findet Regression in einer stark empirisch orientierten Welt in vielfacher Weise Anwendung außerhalb der Mathematik.

Die erhebliche Anwendungsorientiertheit der Regression gibt aber auch Veranlassung und stellt die Notwendigkeit dar das Thema Regression in einem umfassenderen Kontext abzuhandeln.

Derzeit wird in den mathematischen Fachschaften unserer Gymnasien und in Kollegentreffen unter der Leitung unserer Fachberater (sog. Sprengelsitzungen) sehr kontrovers und heftig diskutiert.
(Ich möchte nur beispielhaft an die Diskussion im Zusammenhang mit der generellen Einführung des GTR, der neuen Lehrpläne und der Verfahrensweisen der Qualitätsfindung erinnern)

Die Extrempositionen lauten:

- unsere Schüler sollen lernen sauber zu argumentieren, sorgfältig zu arbeiten und logisch zu denken; dafür böte sich gerade die Mathematik mit ihren Fragestellungen an;
  die in der Schule nur bearbeitbaren Scheinanwendungen würden lediglich unnötig Unterrichtszeit verschwenden.
- Die wenigsten Schüler werden Mathematik in der von uns vermittelten Weise in ihrem späteren Studium oder gar Leben brauchen; um Schüler stärker zu motivieren und die Öffentlichkeit von der zweifelsohne bestehenden Bedeutung von Mathematik zu überzeugen, muss die Nützlichkeit der Mathematik Leitkriterium für unseren Unterricht sein.

Ich möchte auf diese zwei Formen, „Gesichter“ der Mathematik zunächst eingehen, weil meine Stellung dazu auch Grundlage für die Art und Weise sein wird, wie ich Regression hier darstellen werde.

Mathematik kann immer in ihren zwei Gesichtern betrachtet werden:

1. Als eine in sich geschlossene Disziplin, deren Ziel es ist eine möglichtst vollständige, widerspruchsfreie Darstellung zu finden.

Ich möchte dieses Gesicht der Mathematik als ihr „ästhetisches“ Gesicht, ihre **ästhetische Form** bezeichnen.

In dieser ästhetische Form stellt sich Mathematik schon in den „Elemente“ des Euklid von Alexandrien (ca. 365- 300 v. Chr.) dar (in Reprint der Ostwalds Klassiker Nr. 235 ff. , Verlag Harri Deutsch, 1997); einem Lehrbuch der Mathematik, das bis ins 19. Jahrhundert in Auszügen als Schulbuch verwendet wurde.

Welche Wertschätzung und Bedeutung diese Form der Mathematik – bis auf den heutigen Tag - hat, zeigt sich u.A. in einem Dialog des Euklid mit einem seiner Schüler, der auf einen Bericht des Johannes Stobaios (5. Jahrhundert) zurückgeht:

> Ein Schüler soll Euklid nach einiger Zeit des Lernens gefragt haben, welchen Nutzen er von der Mathematik habe. Darauf habe Euklid einem Sklaven befohlen, diesem Schüler eine kleine Geldsumme auszuhändigen, da dieser offenbar bedürftig sei, wenn er die Mathematik um des Nutzens willen betreibe.
> (aus H. Wussing, W. Arnold, Biographien bedeutender Mathematiker, Verlag Volk und Wissen, 1989)

Einen – wenn man so will – nicht sehr zufriedenstellenden Befund diese Form von Mathematik im Hinblick auf ihre Zielsetzung kann man in dem von Kurt Gödel (1906-1978) formulierten und mit formallogischen Mitteln bewiesenen Satz sehen:

„Jedes formale System, das stark genug ist die Arithmetik zu umfassen ist entweder unvollständig (d.h. es gibt modelltheoretisch wahre Sätze, die nicht formal herleitbar sind) oder nicht widerspruchsfrei (d.h. es gibt Sätze die herleitbar sind, deren Negation aber auch herleitbar ist)!“
Insofern unbefriedigend, weil in dem letztgenannten Fall (Gültigkeit von A und von nichtA für irgend ein bestimmtes A) die Herleitbar aller Sätze B möglich ist:
Aus der allg. Gültigkeit von ((A und nichtA) folgt A) und der allg. Gültigkeit von ((A und nichtA) folgt nichtA) stets die Gültigkeit von ((A und nichtA) folgt B) abgeleitet werden kann; und wegen der vorausgesetzten Gültigkeit von (A

und nicht A) auch die allgemeine Gültigkeit von B (gem. Modus Ponens) – und das für alle(!) formalen Sätze B.

Ich meine behaupten zu können, dass diese ästhetische Form der Mathematik, diese in sich geschlossene, in sich logische Form der Mathematik auch heute noch wesentlicher Bestandteil des insbesondere gymnasialen Unterrichts ist.

Das ist auch weiter nicht verwunderlich:

Diese Form ist in den meisten Vorlesungen des Faches Mathematik an den Universitäten, vorzüglich gerade in den Grundvorlesungen gebräuchlich und damit Grundlage für die zukünftigen Lehrer.

Ich möchte an dieser Stelle nicht missverstanden werden:

Ich schätze diese Form Mathematik zu betreiben sehr aber eben als eine der beiden Gesichter der Mathematik!

Für viele von uns war insbesondere die ästhetische Form der Mathematik, wie sie regelmäßig während unserer Schulzeit üblich war, Motivation später Mathematik zu studieren; gab dieses Fach doch eine Sicherheit und (vermeintliche) Gewissheit, die den meisten anderen Fächern fehlt.
Dass wir einer vermeintliche Gewissheit aufgesessen waren, merkten wir oft erst in den ersten Studiensemestern.

Ich möchte einen weiteren Aspekt ansprechen, der diese Form der Mathematik begleitet und wesentlicher Anspruch gymnasialer Erziehung und Bildung ist:

Voraussetzung diese Form von Mathematik zu betreiben ist:

Sauberkeit in der Darstellung, Sorgfalt in der Formulierung und ein hinreichendes Maß an Hartnäckigkeit und Durchhaltevermögen.
Disziplin im geistigen Bereich!
Tugenden, die zu allen Zeiten die Entwicklung der menschlichen Erkenntnis gefördert haben.
Seltsamerweise werden gerade diese Tugenden, welche sich in dieser Form Mathematik zu betreiben am klarsten äußern, von vielen, welche die Qualität des Geistes eher in der Intuition sehen, kritisiert. Sie vergessen dabei, dass der geniale intuitive Einfall oft nur Abschluss langwieriger, disziplinierter Arbeit war!

In diesem Zusammenhang eine Anekdote, die oft zitiert und langzurückgeht:

Von Euklid wird berichtet, dass er seinen König (Ptolemaios I) belehrt habe, dass es keinen einfachen, keinen Königsweg zur Geometrie gäbe (Lexikon des Hellenismus, Herausgeber H. Schmitt, E. Vogt, Verlag Harrassowitz, 2005; hier unter dem Stichwort: Mathematik).

Aus dieser Sichtweise ist die Forderung nach „mehr Spaß im Mathematikunterricht" – wenn man unter „Spaß" momentanes Vergnügen oder Zeitvertreib versteht – eine fragliches Unterfangen und gewinnt erst unter der Forderung nach „mehr Freude an der Mathematik" – in der Bedeutung einer Zufriedenheit nach durchaus mühsamer Suche – ihren Sinn.

Dennoch:

Wir müssen uns klar darüber sein, dass diese ästhetische Form der Mathematik schon immer nur einen kleinen Teil unserer Schüler angesprochen hat und - daraus resultierend - Motivation zum Durchhalten gab, und dass der verständliche Wunsch nach einer vertretbaren Note, einem guten Abitur i.d.R. im Vordergrund steht.

Das dürfte in zunehmendem Maße für unsere heutige Jugend gelten.

Andererseits scheint unsere heutige Jugend sehr viel <u>pragmatischer</u> orientiert zu sein:

Nicht die Kenntnis der inneren Funktionsweise eines Gerätes, sondern die Funktion im Hinblick auf das, was man damit anfangen kann steht im Vordergrund.
Für einen für sich selbst <u>erkannten</u> Nutzen scheint sie bereit zu sein sich erheblicher Mühen zu unterziehen.

Diese Gedanke hat wohl <u>auch</u> dazu beigetragen:

Im gymnasialen Unterricht aber auch in der Abiturprüfung von Baden-Württemberg ab Mitte der 80-er Jahre sogenannten „anwendungsorientierte Aufgaben" einzubeziehen;
und diesbezügliche Inhalte zunehmend in den daran anschließenden Lehrplänen zu formulieren – zugegebenermaßen gerade zu Anfang mit geringem und bis auf den heutigen Tag teilweise fraglichen Erfolg:
Wurde doch diese <u>andere Form</u> Mathematik zu vermitteln von vielen Kollegen als „Verrat" an der gemeinhin vertretenen (ästhetisch) Form der Mathematik angesehen, als Erschwernis im Unterricht.
Zudem wird bei den Schülern die „Nützlichkeit" dieser Form der Aufgabenstellungen häufig in Frage gestellt, weil weder eine Anwendung in

den anderen Fächern noch in ihrer tatsächlichen Lebenswelt für sie erkennbar ist.
Zudem dürften erhebliche Bedenken bestehen, ob wir im schulischen Bereich eine Mathematik vermitteln können, die unseren Schülern einen unmittelbaren Nutzen für ihre aktuelle Lebenswelt liefern kann und
ob diese lediglich an dem aktuellen Nutzen unserer Schüler orientierte Zielsetzung überhaupt mit unserer Aufgabe einer „vertieften Allgemeinbildung" und eine „allgemeinen Studierfähigkeit" in Einklang zu bringen ist.

Realistischer scheint mir zu sein den Schülern zu zeigen, dass Mathematik in vielfacher Weise und in ihren verschiedenen mathematischen Fachmethoden Grundlage und Problemlösungshilfe für andere Unterrichts- und Studienfächer ist.

Fälschlich wird diese Form der Mathematik von Kollegen anderer Fächer, aber auch von uns Mathematikern abschätzig als Hilfswissenschaft bezeichnet, als Magd, die den Fachherrinnen zu dienen habe.
Wohl haben alle Fächer ihre eigene Methodik. Insofern ist die Mathematik hier Magd --- aber eine Magd, welche das Licht der Herrin voranträgt und nicht die Schleppe der Herrin trägt!

Diese ganz andere Deutungsform der Mathematik, die sich eben nicht abschließt, sondern sich in ihrer Symbolik, Axiomatik und Begriffsbildung an außermathematischen, realen Situationen orientiert und
ihnen – und das ist der springende Punkt - eine logische Struktur bieten will, hat viele Namen, die in ihrer Bedeutung nicht immer deckungsgleich sind:

Sie nennt sich „anwendungsbezogen" aber auch „problemlösend",
„technisch/sachlich orientiert", „modellierend" und - mit einem etwas negativen Zungenschlag behaftet – eben auch „mathematische Hilfswissenschaft".

Zur Unterscheidung zu der ästhetischen Form der Mathematik, die weitgehend auf sich selbst bezogen ist, möchte ich dieses zweite Gesicht der Mathematik:

2. Als eine Disziplin ansehen, deren Ziel es ist reale Situationen durch eine je geeignete Form der Symbolik, Axiomatik und Begriffsbildung möglichst nahe zu rekonstruieren um damit die je reale Problemstellungen in ein Problem einer geeigneten mathematischen Fachmethode umzuwandeln, von der man sich eine Lösung der Problemstellung erwartet.

Ich möchte diese zweite Gesicht der Mathematik als „realitätsbezogenes" Gesicht, als ihre **realitätsbezogene Form** bezeichnen.

Da die ästhetische Form der Mathematik gemeinhin bei Mathematikern in Schule und Universität i.d.R. im Vordergrund steht, ist es eher erstaunlich, dass die meisten mathematischen Fachdisziplinen mit ihren je spezifischen Fachmethoden ihre Grundlagen aus realen Fragestellungen abgeleitet haben:

Der grundlegenden Form des menschlichen Zählens und Ordnens stehen die abstrakten Zahlen und Ordnungsrelationen gegenüber.
Jakob Bernoulli (1654-1708) hat Probleme im Zusammenhang mit Glücksspielen erst hinreichend lösen können, nachdem er eine adäquate mathematische Begriffsbildung – die Kombinatorik - für diese reale Problemstellung entwickelt hat (zur „Ars conjectandi" siehe: „Biographien bedeutender Mathematiker", o.g., S 227 ff.).
Für Leibniz und Newton stellte sich das durchaus reale Problem eine geeignete mathematische Begriffsbildung für Fragestellungen wie „Flächenbestimmung krummliniger Flächen" und „momentane Eigenschaften nicht proportionaler Funktionen" zu finden.

Ich möchte daraus sogar die Behauptung wagen, dass
Die meisten mathematischen Disziplinen ihre Grundfragestellung aus realen Situationen ableiten und
viele reale Probleme erst dann lösbar waren weil eine adäquate symbolische, in sich logisch geschlossene – somit eine mathematische – Form gefunden werden konnte.

Ernst Cassirer (1874-1945) geht sogar noch einen Schritt weiter:
Er vermutet, dass für menschliches Denken charakteristisch ist
Symboliken als Bilder von „realen" Gegenständen entwickeln zu können, die keine Ähnlichkeit mit den (realen) Gegenständen haben, sondern zwischen Bild und Gegenstand nur einen logischen Zusammenhang herstellen (Philosophie der symbolischen Formen, E. Cassirer, 1923 aus Gesammelte Werke Band 11, Verlag Felix Meiner, 2001, S. 3 ff.).

In den Anfängen der modernen Physik findet sich bei Galileo Galilei (1564-1642) der Vergleich, dass das „Buch der Natur" in mathematischer Sprache verfasst und nur in mathematischen Chiffren lesbar sei (aus Cassirer, s.o., S. 15)

Unabhängig davon ob man nun an eine „deterministische Natur" glaubt oder eher nach David Hume daran glaubt, dass „Kausalität" nur eine Gewohnheit des menschlichen Geistes ist (Eine Untersuchung über den menschlichen Verstand, D. Hume, 1758 aus dem gleichnamigen Heft 5489, Reclam, 1967) ist ersichtlich, dass **viele reale Probleme erst durch eine je geeignete realitätsbezogene Mathematik zu lösen sind und waren**.

In allen diesen Fällen hat sich die realitätsbezogenen Mathematik nicht als „Schleppenträger" sondern als „Licht", das der Sache den Weg zeigt, erwiesen.

Es ist ein wesentlicher Verdienst von Heinrich Hertz (1857-1894; 1880 Assistent bei H. v. Helmholtz; 1885 bis 1889 Professor an der TH Karlsruhe) und Herrmann v. Helmholtz (1821-1894; zeitweise Professor an der Universität Heidelberg) den Zusammenhang zwischen mathematischer Form und realer Situation näher untersucht zu haben (H. Hertz, Prinzipien der Mechanik in neuem Zusammenhang dargestellt, Gesammelte Werke, 1894; H. v. Helmholtz, „Theorie der Zeichen" aus „Untersuchungen über die Grundlagen der Mathematik und Mechanik").

Nach H. Hertz bestehe das Verfahren der Ableitung darin, dass wir uns „innere Scheinbilder oder Symbole" der äußeren Gegenstände machen, die von solcher Art sind, dass die denknotwendigen Folgen der Bilder stets wieder die Bilder seien von den naturnotwendigen Folgen der abgebildeten Gegenstände.

Dieser Zusammenhang zwischen naturnotwendigen Folgen und denknotwendigen Folgen wird vereinfachend aber anschaulicher durch das teilweise heute noch so genannte Helmholtz-Diagramm (Folie 1) dargestellt, das mit der heute gebräuchlichen „Modellierungs"vorstellung weitgehend übereinstimmt.

„1" stellt die „reale" Situation dar.

„2" die im Hinblick auf „1" „geeignete" symbolische Form, das mathematische Modell, die mathematische Fachmethode, Fachdisziplin etc.

„3" und „4" die jeweils denknotwendigen bzw. naturnotwendigen Folgen aus „2" und „1".

Die Verbindung von „3" und „4" wird häufig als „Korrekturschritt" bezeichnet; sei es dass die Modellierung der individuellen realen Situation zu allgemein ist oder sich gar als ungeeignet erweist.

Über die <u>Notwendigkeit diesen Korrekturschritt zu durchlaufen</u> bestehen sehr verschiedene Vorstellungen.

Für die nur auf Erfahrung und Gewohnheit basierende menschliche Erkenntnis von David Hume wäre das Helmholtz-Diagramm lediglich ein heuristische Verfahren – quasi eine Methode um zu Aussagen zu gelangen, die aber immer wieder erst durch Erfahrung bestätigt werden müssen.

Diametral hierzu die Vorstellung von Rudolf Carnap (1891-1970), der in seiner „Konstitutionstheorie“ jede Erfahrung, jedes „Elementarerlebnis“ - selbst des geisteswissenschaftlichen Bereichs – (zunächst) als in hinreichender Form logistisch darstellbar vermutete (R. Carnap, Der logische Aufbau der Welt, 1928 in gleicher Titel, Verlag Felix Meiner, Band 514, 1998).
Reale Situationen sind hierbei quasi schon und <u>nur</u> durch ihre Begriffsbildung definiert.
Die Verbindung zwischen „1“ und „2“ ist ist hierbei durch eine algorithmische Vorgehensweise definiert; der Korrekturschritt von „3“ zu „4 “ ist obsolet.
Diese extreme Vorstellung basiert auf Überlegungen, die im wesentlich innermathematischen Bereich von Gottlieb Frege und Whitehead/Russell entwickelt wurden.

Unabhängig vor der mehr an Hume oder an Carnap orientierten eigenen Überzeugung erscheint der Korrekturschritt im gymnasialen Unterricht zumindest sehr sinnvoll. Daher habe ich ihn auch in diesem Zusammenhang etwas ausführlicher besprochen:

Geht man davon aus, dass heutige Schüler ihre Motivation für Lerngegenstände vermehrt aus deren „Nützlichkeit“ ziehen, werden sie aus dem Korrekturschritt eine stete Bestätigung finden, dass viele Probleme durch eine realitätsbezogenen Form der Mathematik lösen lassen.

Ich will diesen „Korrekturschritt“ des Helmholtz -Diagramms hier nicht weiter betrachten.

Zentraler scheint mir das Problem der Mathematisierung im Helmholtz-Diagramm:
Für eine je realen Situation ist eine „geeignete“ mathematische Form zu „suchen“.
Ich will zukünftig diese mathematische Form als „(mathematische) Fachmethode“ bezeichnen.

Gehen wir beispielsweise von einem Unterricht aus, der sowohl ästhetische als auch realitätsbezogene Formen der Mathematik berücksichtigt und wohl den „Normalfall“ gymnasialen Mathematikunterrichts darstellt:

a. Anhand eines - oder mehrerer - mehr oder weniger realen Beispiele wird die Notwendigkeit der mathematischen Begriffsbildung gezeigt;
b. die daraus entwickelte Fachmethode wird verallgemeinert und aufgrund von Plausibilitätsbetrachtungen oder Beweisen durch Sätze angereichert;
c. der innermathematische Teil der Fachmethode wird eingeübt;

d. der Unterricht wird durch ein oder mehrere anwendungsbezogene Beispiele, welche mit der erarbeiteten Fachmethode lösbar sind angereichert.

Ich will einen Unterricht der dem o.g. ähnelt durchaus nicht absprechen, dass er für den Schüler förderlich sein kann:

Der Schüler wird mit einem für ihn i.d.R. gut überschaubaren Teilbereich der Mathematik konfrontiert; es wird für ihn klar, welches Wissen erwartbar ist; er kann durch geeignete Übungen das erwartbare Wissen zum Bestandswissen erheben; er erkennt anhand der Beispiele, welche realen Situationen er mit dieser Fachmethode lösen kann.

Gehen wir allerdings davon aus, dass nicht nur die sachliche Forderung nach vertiefter Allgemeinbildung und allgemeiner Studierfähigkeit, sondern auch die aus der Motivation unserer Schüler gestellte Forderung nach Nutzanwendung eine Verstärkung einer realitätsbezogenen Form der Mathematik geradezu zu fordern ist, dann ergeben sich zu dieser Form des Unterricht einige kritische Fragen:

a. In einer realen Situation im späteren Studium oder
während der gymnasialen Schulezeit in anderen Fächern, aber auch im Fach Mathematik selbst, wird dem Studenten oder dem Schüler ja gerade nicht der Hinweis auf die geeignete Fachmethode gegeben:
Er wird sich eventuell an ein ähnliches (zufällig unterrichtetes) Beispiel der Schulzeit oder auch an sehr allgemeine heuristische Hinweise (Vereinfachung des Problems, Abstrahierung,...) erinnern;
i.d.R. wird seine Vorgehensweise eher intuitiv sein;
eventuell wird er mehr oder weniger wahllos einige Fachmethoden, an die er sich noch erinnert, ausprobieren;
die Kenntnis, welche Fachmethoden in der bestimmten Situation nicht zum Ziel führen dürften, und welche eine Chance bieten zum Ziel zu führen, fehlt ihm i.d.R. ganz.
Wie also löst er diese Such- und Abgrenzungsproblem geeigneter Mathematisierung?
b. Um systematisch nach geeigneten Fachmethoden suchen zu können bedarf es eines immer wieder zu aktualisierenden Überblicks über die ihm bekannten Fachmethoden:
In einem Unterricht, der lediglich die aktuelle Fachmethode im Blick hat werden andere Fachmethoden schnell vergessen.
Wie also bleibt ihm ein bewusster Überblick über die einmal erlernten Fachmethoden erhalten?

Geht man davon aus, dass Wiederholen und Vergleichen Grundlage von Lernen ist, dann muss man von einem Unterricht der langfristig die realitätsbezogene Form der Mathematik als wesentlich erachtet fordern, dass

im Unterricht quasi der Ernstfall immer wieder besprochen, diskutiert, geübt wird.

Im Hinblick auf die systematische Suche nach geeigneten Fachmethoden für eine je reale Situation ist daher zu fordern:

a. Bei der Einführung einer neuen Fachmethode anhand von Beispielen ist daher zunächst gemeinsam zu klären,
   welche dieser Beispiele durch bereits bekannte Fachmethoden zu lösen sind/welche nicht, und
   welche Bedingungen an Realbeispiele und an die neue Fachmethode zu knüpfen sind, damit zu erwarten ist, dass die neue Fachmethode eine Lösung bestimmter Realbeispiele verspricht;
   der Schüler muss erkennen, dass eine bestimmte Fachmethode in dieser besonderen Weise eingeführt wurde, um bestimmte Realbeispiel lösen zu können
b. Bei sogenannten Übungsbeispielen (seien sie nun innermathematische oder außermathematisch) ist darauf zu achten, dass
   sie überfachliche Breite haben,
   bewusst die Frage im Vordergrund steht, warum gerade diese oder eine andere Fachmethode Anwendung finden sollte (Bedingungsprüfung!) und
   in geeigneter Weise Realsituationen eingestreut werden, die nicht mit dieser neuen Fachmethode zu lösen sind – sondern ev. mit einer anderen.

Welcher **Grundgedanke** ist Basis **mathematischen Fachmethode „Regression“** ?

Vorgegeben ist ein Satz von Zahlenpaaren (x,y), die i.d.R. durch Messungen vorliegen.

Gesucht ist ein funktionaler Zusammenhang f zwischen den Werten der Zahlenpaare, welcher

1. hinreichend „glatt“ ist und
2. für den alle y-Werte sich „**insgesamt** möglichst wenig“ von den jeweiligen f(x)-Werten „unterscheiden“

(Folie 2)

Greifen wir den vorherigen Gedankengang über die Einführung einer neuen Fachmethode auf, so sollten wir uns Gedanken machen,

1. welche Realsituationen wir mit dieser Fachmethode beschreiben können/welche nicht,
2. wie die zunächst ungenauen Anforderungen an die neue Fachmethode, aber auch an die hierdurch zu mathematisierenden Realsituationen zu präzisieren ist, und
3. welche Präzisierungen wir vornehmen müssen, um die Vorstellung „glatt“ und „insgesamt möglichst wenig unterscheiden“ in mathematisch vertretbarer Form darstellen zu können.

(Folie 3)

Wie müssen wir uns eine <u>Realsituation</u> vorstellen, die durch Regression mathematisierbar ist?

- Das System muss **messbare Größen** (reelle Zahlwerte) als **Eingangswerte** (x-Werte) und **Ausgangswerte** (y-Werte) haben (zur Problematik der „Messbarkeit von Größen" siehe Jaques Hadamard (1865-1963) in Pierre Duhem, Ziel und Struktur der physikalischen Theorien, 1904, in Band 477, Verlag Felix Meiner, 1998, S. 182),
- ein **funktionaler Zusammenhang** im System ist uns **nicht bekannt** (eventuell sind allgemeine Eigenschaften des Systems bekannt),
- das System darf <u>nich</u>t empirisch sein, sondern muss **deterministisch** (kausal, „eindeutig") sein,
- wird ein bestimmter **Eingangsmesswerte mehrfach** gemessen und hat jeweils **verschiedene Ausgangsmesswerte**, so ist dies mit einem **deterministischen System vereinbar**, da wir in der Realität nie ausschließen können, dass wir <u>alle</u> Einflussgrößen bei der Betrachtung des Systems in Rechnung gestellt haben.
- das **System** darf sich **nicht** im Laufe der **Messungen ändern** (kein „Gedächtnis"),
- bei **„gleichem" System** muss erwartet werden, dass sich auch bei **späteren Messungen** y-Werte und f(x)-Werte nicht erheblich unterscheiden,
- es muss eine „hinreichend" **große Zahl an Messwertpaaren** vorhanden sein; die x-Werte sollten den Eingangsmessbereich „dicht" ausfüllen, da ansonsten nicht zu erwarten ist, dass die Regressionsfunktion die y-Werte späterer Messungen gut annähert.

Hinweis:
Die hier gebrachten Einschränkungen an die durch Regression mathematisierbaren Realsituationen und damit auch die mathematischen Anforderungen an die Regression selbst stellen eine eher an unsere **schulischen Möglichkeiten** orientierte Auswahl dar.
Beispielsweise ist auch eine nicht-deterministische Form der Regression als „wahrscheinlichkeitstheoretische Regression" durchaus sinnvoll und wird z.B. in Wikipedia unter dem Stichwort „Regressionsanalyse" näher erläutert.

Betrachten wir nun die Anforderungen, die wir an die Fachmethode „Regression“ stellen müssen, um die o.g. Systeme mathematisieren zu können.

I. Es muss eine (hinreichend große) Zahl an **reellen** Zahlenpaaren (x,y) **fest** vorgegeben sein; hierunter können durchaus Zahlenpaare mit gleichem x-Wert und verschiedenem y-Wert vorkommen aber sogar auch gleiche Paare;
formal werden die Werte als „feste“ Variable mit Index bezeichnet; der Index soll jeweils die Nummer der erfolgten „Messung“ darstellen (z.B. i = 1, 2, ..., n); die Reihenfolge der „Messungen“ soll unerheblich sein.

II. Unter **Vorgabe** eines zunächst recht **einfachen**, „glatten“ **Funktionstyps** (lineare Funktion) ist zu klären, wie sich die Forderung:
Ein bestimmtes f des vorgegebenen Typs muss so auswählbar sein, dass sich **y-Werte und f(x)-Werte „insgesamt möglicht wenig unterscheiden“**
mathematisch verwirklichen lässt.

III. Es wird die Frage geklärt, ob **mit diesem einfachen Typ** einer „linearen“ Regression **auch andere Funktionsvorgaben** behandelt werden können.

IV. Da prinzipiell die Zahl der Messwerte beliebig ergänzt (vergrößert) werden kann, und verschiedene Funktionstypen vorgegeben werden können, ist zu fragen, ob ein **vergleichender Maßstab** gefunden werden kann, der die „Güte“ der jeweils gefundenen Regressionsfunktionen untereinander vergleichbar macht.

V. Da schon vorher darauf hingewiesen wurde, dass ein System nicht nur von einem Eingangswert, sondern von mehreren abhängig sein kann, wird das **Regressionsverfahren auf Messungen erweitert, die je Messung mehr als einen Eingangswert** haben.
Dieses Verfahren wird auch eine Möglichkeit bieten z.B. **polynomiale Funktionstypen** regressiv zu behandeln.

VI. Ein Grundproblem der Behandlung von geeigneten Systemen durch Regression ist, dass man **niemals sicher** sein kann, dass nicht noch ein anderer, hinreichend „glatter“ Funktionstyp existiert, der die **Messwerte besser nähert** oder aber verschiedene, **die ähnlich gut** nähern. Ein Entscheidungsverfahren wird angedeutet.

VII. Es wird auf ein Verfahren hingewiesen **außergewöhnliche Funktionstypen** (z.B. logistische Funktionstypen) in das Regressionsverfahren durch Iteration einzubeziehen.

II Vorgabe eines einfachen („linearen") Funktionstyps
Gesucht ist eine Funktion f dieses Typs, für die sich y_i und $f(x_i)$ „insgesamt möglichst wenig unterscheiden"
Gesucht ist eine geeignete Mathematisierung des Begriffs „insgesamt möglichst wenig unterscheiden"

Variante 1:

$$\left.\begin{aligned} f(x) &= \beta_0 + \beta_1 x \\ y &= \beta_0 + \beta_1 x \end{aligned}\right\} \beta_0, \beta_1 \in \mathbb{R}$$

gesucht: $f(x) = \hat{\beta}_0 + \hat{\beta}_1 x$

mit „Minimalanforderung"

$\wedge$ – Variablennamen immer im Hinblick auf feste (x_i, y_i) Zahlenwerte

Mathematisierungsvariante 1:

$$\sum_{i=1}^{n} u_i = \sum (y_i - f(x_i)) =: S_1(\beta_0, \beta_1)$$

Summe der Unterschiede

Ziel: $\hat{\beta}_0, \hat{\beta}_1$ so, dass $S_1(\hat{\beta}_0, \hat{\beta}_1)$ minimal!

$$\begin{aligned} S_1(\beta_0, \beta_1) &= \sum (y_i - f(x_i)) \\ &= \sum (y_i - \beta_0 - \beta_1 x_i) \end{aligned}$$

$$S_1(\beta_0, \beta_1) = \sum y_i - n\beta_0 - \beta_1 \sum x_i$$

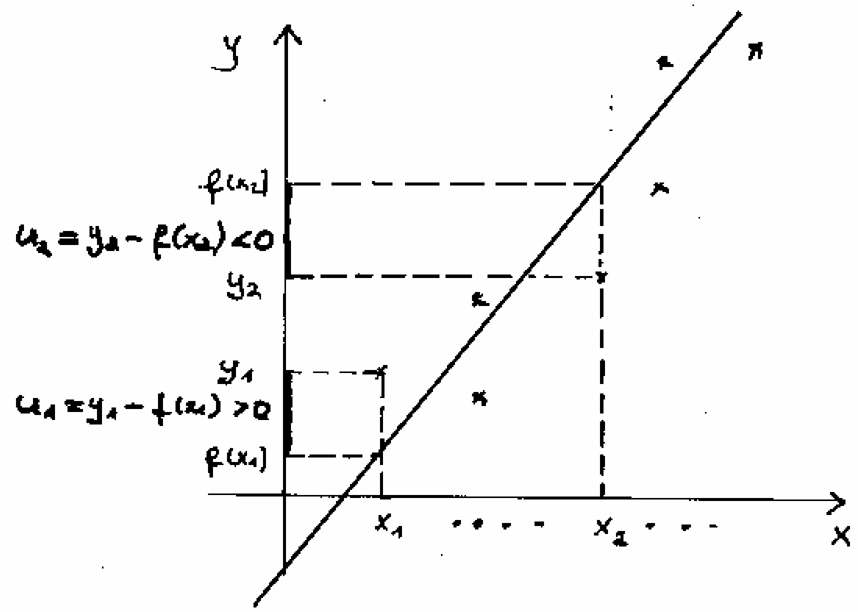
$\sum y_i, \sum x_i, n$ feste Zahlwerte!

Gleichung mit den drei Variablen $S_1(\beta_0, \beta_1), \beta_0, \beta_1$
$\longrightarrow$ Lösungsmethode: Analytische Geometrie!

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ S_1(\beta_0, \beta_1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sum y_i \end{pmatrix} + \beta_0 \begin{pmatrix} 1 \\ 0 \\ -n \end{pmatrix} + \beta_1 \begin{pmatrix} 0 \\ 1 \\ -\sum x_i \end{pmatrix}$$

„schräge" Ebene; kein Minimum für $S_1(\beta_0, \beta_1)$ vorhanden
$\longrightarrow$ Mathematisierungsvariante 1 ungeeignet

– 14 –


y
f(x2)
u2 = y2 - f(x2) < 0
y2
y1
u1 = y1 - f(x1) > 0
f(x1)
x1
x2
X

Eventuell ist Variante 1 ungeeignet, weil sich die u_i „gegenseitig aufheben" (u_i teilweise negativ, teilweise positiv!)

Variante 2:

Mathematisierungsvariante 2:

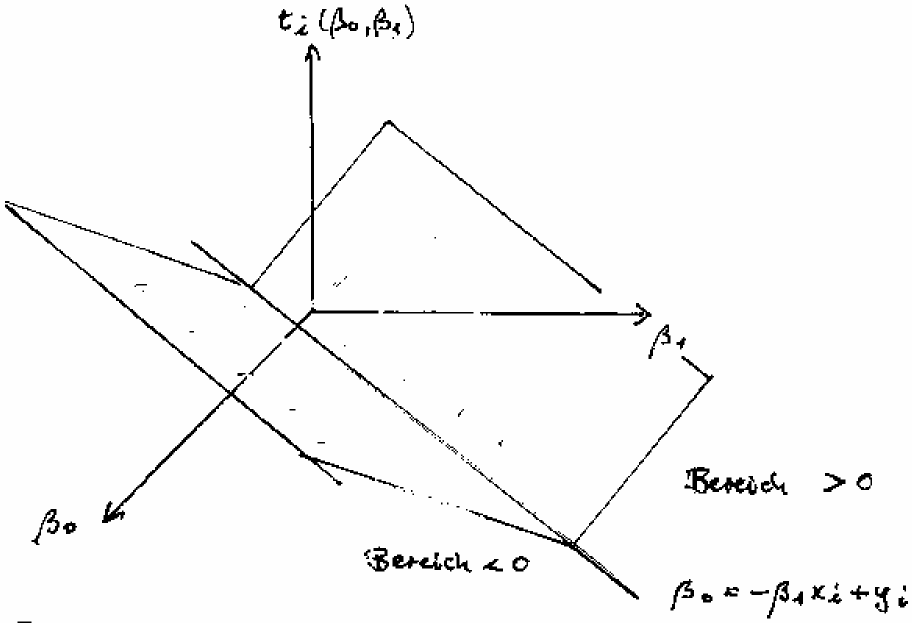
$$S_2(\beta_0,\beta_1) = \sum u_i = \sum |y_i - f(x_i)| \geqq 0$$

Betrachtet man hierzu einen Teilsummand von S_2:

$$t_i(\beta_0,\beta_1) := |y_i - f(x_i)|$$
$$= |y_i - \beta_0 - \beta_1 x_i|$$

$$t_i(\beta_0,\beta_1) = \begin{cases} y_i - \beta_0 - \beta_1 x_i & \text{falls } y_i - \beta_0 - \beta_1 x_i > 0 \\ 0 & \text{falls } y_i - \beta_0 - \beta_1 x_i = 0 \\ & (\Leftrightarrow \beta_0 = -\beta_1 x_i + y_i) \\ -y_i + \beta_0 + \beta_1 x_i & \text{falls } y_i - \beta_0 - \beta_1 x_i < 0 \end{cases}$$

Darstellung wieder in β_0, β_1, $t_i(\beta_0,\beta_1)$ – Form



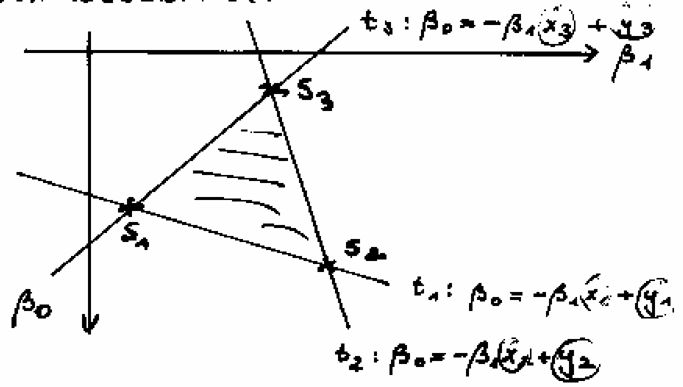
Betrachtet man $S_2(\beta_0,\beta_1)$ insgesamt in der gleichen Darstellung, so erhält man ein Gebilde, das keine Ebene ist aber

a) immer „oberhalb" der β_0,β_1–Ebene liegt (klar: $S_2(\beta_0,\beta_1) \geq 0$)
b) aus ebenen Vielecken zusammengesetzt ist
warum?

— 16 —

Am Beispiel $n = 3$:

Nur der Verlauf von t_1, t_2, t_3 in der β_0/β_1-Ebene ist wesentlich:



Innerhalb jeden, durch die Geraden eingegrenzten Bereichs lässt sich $S_2(\beta_0, \beta_1)$ betragsfrei darstellen, da die $y_i - f(x_i)$ keinen Vorzeichenwechsel haben.
Innerhalb jeder dieser Bereiche lässt sich daher Variante 1 anwenden (Teilausschnitt aus einer Ebene)

Es kann nun gezeigt werden, dass $S_2(\beta_0, \beta_1)$ höchstens an den Schnittpunkten S_1, S_2, S_3 minimal wird.

Beweis:

a) „elementar": Durch Betrachtung der „schrägen" Vielecke

b) Mit Hilfe der Fachmethode „Lineares Optimieren"

(z.B. in: K. Neumann, Operations Research – Verfahren, Band I, Verlag Carl Hauser, 1975, S. 18ff.)

→ Mathematisierungsvariante 2 geeignet

Diese Mathematisierungsvariante findet allerdings nur zögernd Anwendung, weil bei einer (wie geforderten) großen Zahl von Messungen der Rechenaufwand erheblich ist.

Eventuell könnte Variante 2 so modifiziert werden, dass sich $S(\beta_0,\beta_1)$ in „geschlossener" Form darstellen lässt und alle Teilsummanden positiv bleiben

Variante 3:

Mathematisierungsvariante 3:

$$S(\beta_0,\beta_1) := \sum u_i^2 = \sum (y_i - f(x_i))^2 \quad \geq 0$$
$$= \sum (y_i - \beta_0 - \beta_1 x_i)^2$$

Diese Variante hat den „Fehler", dass für $|u_i| < 1 \to u_i^2 < |u_i|$ und $|u_i| > 1 \to u_i^2 > |u_i|$
Dieser „Fehler" wird allerdings auch z.B. für die gesamte Wahrscheinlichkeitsrechnung akzeptiert:
$\text{Varianz}(X) = E((X-E(X))^2)$

Da es sich nun bei $S(\beta_0,\beta_1)$ um eine hinreichend oft differenzierbare Funktion von β_0,β_1 handelt, wird man zur Bestimmung des Minimums die Fachmethode der Differentialrechnung in Anwendung bringen können.

$$\frac{\partial S}{\partial \beta_0}(\beta_0,\beta_1) = \sum 2\cdot(y_i-\beta_0-\beta_1 x_i)(-1)$$
$$= (-2)\sum(y_i-\beta_0-\beta_1 x_i)$$

$$\frac{\partial S}{\partial \beta_0}(\hat\beta_0,\hat\beta_1) = 0 \quad (\text{notwendige Bedingung für Minimum})$$

$$(-2)\sum(y_i-\hat\beta_0-\hat\beta_1 x_i) = 0$$
$$\sum y_i - n\hat\beta_0 - \hat\beta_1 \sum x_i = 0$$
$$(1)\quad n\cdot\hat\beta_0 + \left(\sum x_i\right)\hat\beta_1 = \sum y_i$$

$$\frac{\partial^2 S}{\partial \beta_0^2}(\beta_0,\beta_1) = (-2)\sum(-1) = 2n > 0$$

$$\frac{\partial^2 S}{\partial \beta_0^2}(\hat\beta_0,\hat\beta_1) > 0 \quad (\text{hinreichende Bedingung für Minimum})$$

— 18 —

<u>analog:</u>

$$\frac{\partial S}{\partial \beta_1}(\beta_0,\beta_1) = \sum 2(y_i-\beta_0-\beta_1 x_i)(-x_i)$$
$$= (-2)\cdot\sum\left[(y_i-\beta_0-\beta_1 x_i)\, x_i\right]$$

$\frac{\partial S}{\partial \beta_1}(\hat\beta_0,\hat\beta_1) = 0$ (notwendige Bedingung für Minimum)

$$(-2)\sum(x_i y_i-\hat\beta_0 x_i-\hat\beta_1 x_i^2) = 0$$
$$\sum(x_i y_i)-\hat\beta_0\sum x_i-\hat\beta_1\sum x_i^2 = 0$$
$$(2)\ (\sum x_i)\hat\beta_0+(\sum x_i^2)\hat\beta_1 = \sum x_i y_i$$

$\frac{\partial^2 S}{\partial \beta_1^2}(\beta_0,\beta_1) = (-2)\sum(-x_i^2) = 2\sum x_i^2 > 0$ falls mindestens ein $x_i \neq 0$

$\frac{\partial^2 S}{\partial \beta_1^2}(\hat\beta_0,\hat\beta_1) > 0$ (hinreichende Bedingung für Minimum)

$$\boxed{\begin{array}{ll}(1) & n\hat\beta_0+(\sum x_i)\hat\beta_1 = \sum y_i\\ (2) & (\sum x_i)\hat\beta_0+(\sum x_i^2)\hat\beta_1 = \sum x_i y_i\end{array}}$$

„Normalgleichungen" der Regression

<u>Lösung:</u>

$$\hat\beta_1 = \frac{\sum x_i y_i-\frac{1}{n}(\sum x_i)(\sum y_i)}{\sum x_i^2-\frac{1}{n}(\sum x_i)^2}$$
$$\hat\beta_0 = \frac{1}{n}\sum y_i-\left(\frac{1}{n}\sum x_i\right)\hat\beta_1$$

$\longrightarrow$ Mathematisierungsvariante 3 <u>geeignet</u>

Allerdings mit den (eher unbedeutenden) Einschränkungen

a) $x_i \neq 0$ für zumindest ein x_i

b) Normalengleichungen Lösbar

und der generellen Problematik u. ②

Zukünftig wird die Variante 3 als Mathematisierung der Fachmethode „Regression" zugrundegelegt.

— 19 —

III Es wird geklärt, ob mit diesem einfachen Typ ein „linearen" Regression auch andere Funktionsvorgaben behandelt werden können

1. Funktionstyp „proportionale" Funktion:

Häufig wird man vermuten, dass die Messwertpaare zueinander proportional sind

Der Funktionstyp hat dann die Form

$$f(x) = \beta_1 x$$
$$y = \beta_1 x$$

Gesucht wird $f(x) = \hat{\beta}_1 x$ mit „Minimalabweichung"

$$S(\beta_1) = \sum (y_i - f(x_i))^2$$
$$= \sum (y_i - \beta_1 x_i)^2$$

$$\frac{\partial S}{\partial \beta_1}(\beta_1) = \sum 2(y_i - \beta_1 x_i)(-x_i)$$
$$= (-2) \sum (x_i y_i - \beta_1 x_i^2)$$

$$\frac{\partial S}{\partial \beta_1}(\hat{\beta}_1) = 0 \quad \text{(notwendige Bedingung für Minimum)}$$
$$(-2) \sum (x_i y_i - \hat{\beta}_1 x_i^2) = 0$$
$$\sum x_i y_i - \hat{\beta}_1 \sum x_i^2 = 0$$
$$\boxed{\hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2}} \quad \text{„Normalgleichung" der Regression}$$

$$\frac{\partial^2 S}{\partial \beta_1^2}(\beta_1) = (-2) \sum (-x_i^2) = (+2) \sum x_i^2 > 0 \quad \text{falls zumindest ein } x_i \neq 0$$
(hinreichende Bedingung für Minimum)

Es wird sich allerdings herausstellen, dass diese Form der Regression einer „proportionalen" Funktion ihre „Tücken" hat!

2. Rückführung von bestimmten Funktionstypen durch Transformation auf eine lineare Funktion

Häufig wird zur Auswertung von Messwertpaaren einfaches oder doppelt-logarithmisches Papier verwendet:
Einfach-logarithmisch: x-Skala linear; y-Skala logarithmisch
Doppelt-logarithmisch: x-Skala logarithmisch; y-Skala logarithmisch

In diesen Anwendungsfällen wird erwartet, dass das Schaubild, das die transformierten Wertepaare annähert von linearem Typ ist.

Vorgehensweise bei diesem Verfahren:

(1) (x_i, y_i) – Messwertpaare

(2) (a_i, b_i) – Wertepaare durch Transformationen
$$a = \tau_1(x)\ ;\ b = \tau_2(y)$$

(3) Lineare Funktion g durch „lineare" Regression der Wertpaare (a_i, b_i)

(4) Rücktransformation von g zu f

a. Einfach-logarithmische Transformation:

Transformation:

$$a = \tau_1(x) = x\ ;\ \ b = \tau_2(y) = \log y$$

Lineare Funktion g durch „lineare" Regression:

$$g(a) = b = \hat{\beta}_0 + \hat{\beta}_1 a$$

Rücktransformation von g zu f:

$$b = \hat{\beta}_0 + \hat{\beta}_1 a\ ;\quad \log y = \hat{\beta}_0 + \hat{\beta}_1 x\ ;$$
$$e^{\log y} = e^{\hat{\beta}_0 + \hat{\beta}_1 x}\ ;\quad y = e^{\hat{\beta}_0} \cdot e^{\hat{\beta}_1 x}$$
$$y = c_1 \cdot e^{c_2 x}$$

Mit dieser Transformation lassen sich Regressionen für den Funktionstyp „exponentielles Wachstum"
$$y = c_1 \cdot e^{c_2 x}$$
darstellen.

b. Doppelt-logarithmische Transformation:

Transformation:

$a = \tau_1(x) = \log x$; $b = \tau_2(y) = \log y$

Lineare Funktion g durch „lineare" Regression:

$g(a) = b = \hat{\beta}_0 + \hat{\beta}_1 \cdot a$

Rücktransformation von g zu f:

$b = \hat{\beta}_0 + \hat{\beta}_1 \cdot a$; $\log y = \hat{\beta}_0 + \hat{\beta}_1 \cdot \log x$;

$e^{\log y} = e^{\hat{\beta}_0 + \hat{\beta}_1 \log x}$; $y = e^{\hat{\beta}_0} \cdot (e^{\log x})^{\hat{\beta}_1}$

$y = e^{\hat{\beta}_0} \cdot x^{\hat{\beta}_1}$; $y = c_1 \cdot x^{c_2}$

Mit dieser Transformation lassen sich Regressionen für den Funktionstyp „Potenzfunktion"

$$y = c_1 \cdot x^{c_2}$$

darstellen.

3. Hinweise:

a. Voraussetzung für das Transformationsverfahren muss die Invertierbarkeit von τ_1, τ_2 sein

b. In CAS-Systemen kann man entweder:
x-y-Listen erstellen, diese zu a-b-Listen transformieren und darauf die lineare Regression anwenden
oder:
man verwendet unmittelbar die zu bestimmten Funktionstypen vorhandenen Regressionsverfahren

c. Vorsicht!
Bisher haben wir nur die transformierten Wertepaare über die „lineare" Regression durch eine lineare Funktion g „angenähert".
Es wurde nicht gezeigt, dass die rücktransformierte Funktion f dann auch diese Eigenschaft hat.

IV Da prinzipiell die Zahl der Messwertpaare beliebig ergänzt (vergrößert) werden kann bzw. verschiedene Funktionstypen vorgegeben werden können, ist zu fragen, ob ein vergleichender Maßstab gefunden werden kann, der die „Güte" der jeweils gefundenen Regressionsfunktionen untereinander vergleichbar macht

Vorgabe: n Messwertpaare (x_i, y_i) $i = 1, \dots n$

Funktionstyp: Lineare Funktion $y = \beta_0 + \beta_1 x$

$\Rightarrow$ Regressionsgerade $y = \hat{\beta}_0 + \hat{\beta}_1 x$

Bezeichnungen: $\hat{y}_i := f(x_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i$ „Prognosewerte"

$\hat{u}_i := y_i - \hat{y}_i$ „Unterschiedswerte" „Restwerte" „Residuen"

Ziel: Gesucht ist ein Maß M, das

(1) „unabhängig" von n ist;

(2) $0 \leq M \leq c$ c fest für alle Messungen

(3) Falls y_i „ideal" auf einer Geraden liegen d. h. $y_i = \hat{y}_i$ dann: $M = c$

(4) Falls die Prognosewerte $\hat{y}_i$ „unabhängig" von x_i sind d. h. $\hat{y}_i = \text{const.}$ dann: $M = 0$

Variante 1:

Wir erinnern uns, dass $\hat\beta_0$ und $\hat\beta_1$ so gewählt wurde, dass

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum u_i^2 \geq 0$$

minimal wird:

$$S(\hat\beta_0, \hat\beta_1) = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - \hat y_i)^2 = \sum_{i=1}^{n} \hat u_i^2 \geq 0$$

Die Frage ist, ob $\sum_{i=1}^{n} u_i^2 =: M$ ein Maß darstellt, das die Bedingungen (1) bis (4) erfüllt!

- Sicher ist $M \geq 0$ damit ein Teil von (2) erfüllt
- Für den „Idealfall": y_i liegen alle auf einer Geraden; somit $y_i = \hat y_i$ gilt:

$$M := \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (y_i - \hat y_i)^2 = 0 = c^*$$

aber:

(Folie 2)

Wenn die o.g. Messung mit n Messwertpaaren und $M(n) = \sum_{i=1}^{n} \hat u_i^2$ um weitere Messwertpaare so ergänzt wird, dass annähernd die gleiche Regressionsgerade entsteht (was zu vermuten ist!) dann gilt:

$$M(n+n') = \sum_{i=1}^{n+n'} \hat u_i^2 = \sum_{i=1}^{n} \hat u_i^2 + \sum_{i=n+1}^{n+n'} \hat u_i^2$$

$$= M(n) + \sum_{i=n+1}^{n+n'} \hat u_i^2$$

Da zu vermuten ist, dass nahezu alle $u_i \neq 0$ bei gleicher Regression und wachsender Messwertezahl, wird M beliebig groß und ist von n abhängig

$\longrightarrow$ Die so gewählte Definition von M ist somit ungeeignet.

Variante 2:

Um M so zu definieren, damit die Bedingungen (1) bis (4) erfüllt sind, muss in die Definition Folgendes eingehen:

— Eine Mittelung über die Zahl „n" der Versuche

$$\bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i ;$$

— die $\hat{u}_i = y_i - \hat{y}_i$ (Beachtung des Idealfalls $y_i = \hat{y}_i$)

— $\hat{y}_i$ im Vergleich zu $\bar{y}$ ($\hat{y}_i$ unabhängig von x_i; $\hat{y}_i = \text{const.}$)

Bezeichnungen:

$V := \sum_i (y_i - \bar{y})^2$ „Gesamtvariabilität" der y_i – Messwerte

$\hat{V} := \sum_i (\hat{y}_i - \bar{y})^2$ "Variabilität der Prognosewerte"

$V_R := \sum_i (y_i - \hat{y}_i)^2 = \sum_i \hat{u}_i^2$ "Variabilität der Residuen"

Analyse in den Extremfällen: „ideal", „unabhängig":

	Idealfall $y_i = \hat{y}_i$	Unabhängigkeitsfall $\hat{y}_i = \text{const.}$
$V = \sum_i (y_i - \bar{y})^2 =$	V	V
$\hat{V} = \sum_i (\hat{y}_i - \bar{y})^2 =$	V	0
$V_R = \sum_i (y_i - \hat{y}_i)^2$	0	V
$\frac{\hat{V}}{V}$	1	0
$\frac{V_R}{V}$	0	1

Es ist zu vermuten:

$$(*) \quad V = \hat{V} + V_R$$

Im Hinblick auf Normierungsgründen damit:

$$1 = \frac{\hat{V}}{V} + \frac{V_R}{V}$$

Es ist sinnvoll zu definieren

$$\boxed{M := \frac{\hat{V}}{V}}$$

denn: M ist durch V „normiert"; $0 \leq M \leq 1$; $M = 1$ im „Idealfall"; $M = 0$ im „Unabhängigkeitsfall"

Bezeichnungen:

$$M = \frac{\hat{V}}{V} =: R^2 \quad \text{„Bestimmtheitsmaß"}$$

$$K := \sqrt{R^2} = |R| \quad \text{„Korrelationskoeffizient"}$$

$$\boxed{K = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}}}$$

Beweis der Gleichung (*) $V = \hat{V} + V_R$:

(*) lässt sich nicht „elementar" beweisen

Gesucht ist also nach einer math. Methode, welche für die Beweisführung geeignet ist.

Wir erinnern uns an die „Normalgleichungen" aus denen wir $\hat{\beta}_0$ und $\hat{\beta}_1$ entnehmen konnten:

$$n \cdot \hat{\beta}_0 + (\Sigma x_i) \cdot \hat{\beta}_1 = \Sigma y_i$$
$$(\Sigma x_i) \cdot \hat{\beta}_0 + (\Sigma x_i^2) \cdot \hat{\beta}_1 = \Sigma x_i y_i \quad \text{„Normalgleichungen"}$$

Es handelt sich offenbar um eine (2×2)-Abbildung:

$$\underset{(2\times 2)}{\begin{pmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}} \underset{(2\times 1)}{\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix}} = \underset{(2\times 1)}{\begin{pmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{pmatrix}}$$

Bezeichnungen:

$$\underset{(2\times 1)}{\vec{\beta} := \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix}} \quad X := \underset{(n\times 2)}{\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \quad \vec{y} := \underset{(n\times 1)}{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}$$

$$X^T X = \underset{(2\times n)}{\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}} \cdot \underset{(n\times 2)}{\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} = \begin{pmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{pmatrix}$$

$$X^T \vec{y} = \underset{(2\times n)}{\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}} \underset{(n\times 1)}{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}} = \underset{(2\times 1)}{\begin{pmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{pmatrix}}$$

Damit: $(X^T X)\vec{\beta} = X^T \vec{y}$ **Normalgleichungen in Vektorform**

Falls $X^T X$ invertierbar: $\hat{\vec{\beta}} = (X^T X)^{-1} X^T y$ (für CAS-System gut lösbar!)

— 26 —

Beweis von (*) $V = \hat{V} + V_R$ mit Hilfe der Vektorrechnung:

Dürch einfaches „Rechnen“ ergibt sich:

$$\underline{V} = \sum_i (y_i - \bar{y})^2 = (y_1 - \bar{y}, \ldots, y_n - \bar{y}) \cdot \begin{pmatrix} y_1 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{pmatrix}$$

$$= (\vec{y} - \bar{y} \cdot 1)^T (\vec{y} - \bar{y} \cdot 1)$$

Hinweis: Die Transposition „T“ ist linear aber „anti kommutativ“: $(A \cdot B)^T = B^T \cdot A^T$

$$1 := \overset{(n \times 1)}{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}$$

$$= \vec{y}^T \vec{y} - \bar{y}(1^T \vec{y}) - \bar{y}(\vec{y}^T 1) + \bar{y}^2 (1^T 1)$$

$$= \vec{y}^T \vec{y} - n\bar{y}^2 - n\bar{y}^2 + n\bar{y}^2$$

$$= \underline{\vec{y}^T \vec{y} - n\bar{y}^2}$$

Hinweis:

$1^T \vec{y} = \sum y_i = n \cdot \bar{y}$

$\vec{y}^T 1 = \sum y_i = n \cdot \bar{y}$

$$\underline{\hat{V}} = \sum_i (\hat{y}_i - \bar{y})^2 = (\hat{\vec{y}} - \bar{y} 1)^T (\hat{\vec{y}} - \bar{y} 1)$$

$$= \hat{\vec{y}}^T \hat{\vec{y}} - \bar{y}(1^T \hat{\vec{y}}) - \bar{y}(\hat{\vec{y}}^T 1) + n\bar{y}^2$$

$$\hat{\vec{y}} := \overset{(n \times 1)}{\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix}}$$

Hinweis (einzig nicht unmittelbarer „Rechenschritt“):

(a) $$X \cdot \hat{\vec{\beta}} = \overset{(n \times 2)}{\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \overset{(2 \times 1)}{\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix}} = \begin{pmatrix} \hat{\beta}_0 + \hat{\beta}_1 x_1 \\ \vdots \\ \hat{\beta}_0 + \hat{\beta}_1 x_n \end{pmatrix} = \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix} = \hat{\vec{y}}$$

(b) $$X^T X \hat{\vec{\beta}} = X^T \vec{y}$$ Normalengleichungen! und wegen (a):

$$X^T \hat{\vec{y}} = X^T \vec{y}$$

$$\overset{(2 \times n)}{\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}} \overset{(n \times 1)}{\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix}} = \overset{(2 \times n)}{\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}} \overset{(n \times 1)}{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}$$

$$\begin{pmatrix} \sum \hat{y}_i \\ \sum x_i \hat{y}_i \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

somit:

$\rightarrow \sum \hat{y}_i = \sum y_i$

$\rightarrow \frac{1}{n} \sum \hat{y}_i = \frac{1}{n} \sum y_i = \bar{y}$ (!)

(**) $\rightarrow \underline{1^T \hat{\vec{y}} = 1^T \vec{y}}$; $\underline{\hat{\vec{y}}^T 1 = \vec{y}^T 1}$

$$= \hat{\vec{y}}^T \hat{\vec{y}} - \bar{y}(1^T \vec{y}) - \bar{y}(\vec{y}^T 1) + n\bar{y}^2$$

$$= \hat{\vec{y}}^T \hat{\vec{y}} - n\bar{y}^2 - n\bar{y}^2 + n\bar{y}^2$$

$$= (X\hat{\vec{\beta}})^T (X\hat{\vec{\beta}}) - n\bar{y}^2$$ wegen (a)

$$= \hat{\vec{\beta}}^T X^T X \hat{\vec{\beta}} - n\bar{y}^2$$ wegen Normalengleichungen

$$= \underline{\hat{\vec{\beta}}^T X^T \vec{y} - n\bar{y}^2}$$

– 27 –

$$
\begin{aligned}
\underline{V_R} &= \sum_i (y_i - \hat{y}_i)^2 = (\vec{y} - \hat{\vec{y}})^T (\vec{y} - \hat{\vec{y}}) \\
&= \vec{y}^T \vec{y} - \hat{\vec{y}}^T \vec{y} - \vec{y}^T \hat{\vec{y}} + \hat{\vec{y}}^T \hat{\vec{y}} \\
&= \vec{y}^T \vec{y} - (X\vec{\beta})^T \vec{y} - \vec{y}^T (X\vec{\beta}) + (X\vec{\beta})^T (X\vec{\beta}) \\
&= \vec{y}^T \vec{y} - \vec{\beta}^T X^T \vec{y} - \vec{y}^T (X\vec{\beta}) + \vec{\beta}^T X^T X \vec{\beta} \\
&= \vec{y}^T \vec{y} - \vec{\beta}^T X^T \vec{y} - \vec{y}^T (X\vec{\beta}) + \vec{\beta}^T X^T \vec{y} \\
&= \vec{y}^T \vec{y} - \vec{y}^T (X\vec{\beta}) \\
&= \underline{\vec{y}^T \vec{y} - \vec{\beta}^T X^T \vec{y}}
\end{aligned}
$$

<u>Hinweis:</u>

$$\hat{\vec{y}} = X\vec{\beta}$$

$$X^T X \vec{\beta} = X^T \vec{y}$$

<u>Hinweis:</u>

$$\vec{y}^T X \vec{\beta} \overset{\text{Zahl!}}{=} (\vec{y}^T X \vec{\beta})^T = \vec{\beta}^T X^T \vec{y}$$

Vergleicht man die Ergebnisse der drei Rechnungen, so erhält man $V = \hat{V} + V_R$

q.e.d.

<u>Hinweis zur sogenannten „proportionalen" Regression</u>

Es würde bereits darauf hingewiesen, dass wohl zu Funktionen des Typs

$$y = \beta_1 x$$

eine Regression $(\, y = f(x) = \hat{\beta}_1 x \,)$ mit

$$\sum x_i^2 \cdot \hat{\beta}_1 = \sum x_i y_i \qquad \text{Normalengleichung}$$

möglich ist, aber ihre „Tücken" hätte!

Das Problem ist, dass der aus der Identität $V = \hat{V} + V_R$ entnommene Korrelationskoeffizient so nicht mehr definierbar ist, weil $V \neq \hat{V} + V_R$ i.A.

Wähl man hier im Hinblick auf die Normalengleichung

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

so erhält man wieder die Normalengleichung

$$X^T X \hat{\beta}_1 = X \vec{y}$$

Weiterhin gilt: $X\hat{\beta}_1 = \hat{\vec{y}}$ und somit $X^T \hat{\vec{y}} = X^T \vec{y}$

$X^T \hat{\vec{y}} = X^T \vec{y} \leftrightarrow \sum x_i \hat{y}_i = \sum x_i y_i$ daraus ist aber <u>nicht</u> $\sum \hat{y}_i = \sum y_i$ zu entnehmen folglich $(x$[illegible]$)$ $1^T \hat{\vec{y}} \neq 1^T \vec{y}$

V. Regressionsverfahren, das je Messung mehr als einen Eingangswert aber weiterhin einen Ausgangswert hat

Vorgabe:

n Messwertpunkte $(x_{1i}, \dots x_{ki}, y_i)$ $i = 1, \dots, n$

Jeder Messwertpunkt besteht somit aus K Eingangswerten und einem Ausgangswert y

Funktionstyp soll sein:

$$y = f(\vec{x}) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

Gesucht sind wieder Zahlen $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$, so dass die Funktion

$$f(\vec{x}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k$$

die Unterschiede $u_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_k x_{ki})$

„möglichst gering" werden lässt.

→ Multilineare Regression

Bedeutung für Realsituationen:

(1) Wir haben bereits festgestellt, dass im elementaren Fall (ein Eingangswert, ein Ausgangswert) durchaus bei zwei Messungen mit gleichem x-Wert verschiedene y-Werte als Ausgangsmessung vorkommen können.

Dies ist im deterministischen Fall nur möglich, wenn neben dem x-Wert der Messung noch ein weiterer (uns zunächst unbekannter) Wert die Ausgangsmessung beeinflusst.

Um Realsituationen mit mehreren Eingangswerten bewältigen zu können, benötigen wir die Multilineare Regression.

(2) Es ist bereit in $\overline{III}$ besprochen worden, dass durch Transformation bestimmte Funktionstypen mit Hilfe des Verfahrens der Linearen Transformation mit dem Regressions verfahren bearbeitet werden können.

Insbesondere der Funktionstyp des nichtlinearen Polynoms kann (wegen der i.A. fehlenden Invertierbarkeit) mit dieser Transformationsmethode nicht bearbeitet werden.

Gehen wir beispielsweise von Messwertpaaren (x_i, y_i) aus und wollen eine Regression vom Typ eine quadratischen Funktion durchführen, so ist dieses Problem gleichwertig zu einem Problem einer Multilinearen Regression mit $k=2$:

Messwertpunkte: (x_{1i}, x_{2i}, y_i) mit $x_{1i} = x_i$
$x_{2i} = x_i^2$

Die daraus resultierende Regressionsfunktion hat dann die Form

$$y = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$$

<u>Wie ist das Regressionsproblem für Multilineare Regressionen zu Lösen?</u>

Ausgehen von den Messwertpunkten $(x_{1i}, \dots x_{ki}, y_i)$ sind $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ zu finden, damit

$$S(\beta_0, \beta_1, \dots, \beta_k) := \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{ki})^2$$

minimal wird

Somit wieder:

$$\frac{\partial S(\beta_0, \beta_1, \dots, \beta_k)}{\partial \beta_0} = 0$$

was zu $n \cdot \hat{\beta}_0 - (\sum x_{1i})\hat{\beta}_1 + \dots + (\sum x_{ki})\hat{\beta}_k = \sum y_i$ führt

$$\frac{\partial S(\beta_0, \beta_1, \dots, \beta_k)}{\partial \beta_j} = 0 \qquad (j = 1, \dots, k)$$

was zu $(\sum x_{ji})\hat{\beta}_0 + (\sum x_{ji} \cdot x_{1i})\hat{\beta}_1 + \dots + (\sum x_{ji} \cdot x_{ki})\hat{\beta}_k = \sum y_i x_{ji}$ führt

Die $k+1$ Gleichungen heißen <u>Normal[redacted]gleichungen</u> der Multilinearform

Die Anwendung der Vektormethode führt
führt nun wieder zu einer drastischen Vereinfachung

Bezeichnungen:

$$\vec{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad X := \begin{pmatrix} 1 & x_{11} & \cdots & x_{k1} \\ \vdots & & & \\ 1 & x_{1n} & \cdots & x_{kn} \end{pmatrix} \qquad \hat{\vec{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_k \end{pmatrix}$$

Wie man mit etwas nachdenken erkennt gilt hiermit

$$X^T X \hat{\vec{\beta}} = X^T \vec{y}$$ Normalengleichungen der Multilinearform in Vektordarstellung

und falls $X^T X$ invertierbar ist

$$\hat{\vec{\beta}} = (X^T X)^{-1} X^T \vec{y}$$ (mit CAS gut darstellbar!)

Insbesondere gelten auch alle Rechenschritte zum Nachweis für die Identität (*) $V = \hat{V} + V_R$.

Somit ist die Definition des (multiplen)

Bestimmtheitsmaßes $R^2 := \frac{\hat{V}}{V} = \frac{\Sigma(\hat{y}_i - \bar{y})^2}{\Sigma(y_i - \bar{y})^2}$

sinnvoll;

und damit auch der (multiple)

Korrelationskoeffizient $\kappa := \sqrt{R^2} = |R|$

— 31 —

VI Hinweise für eine geeignete Auswahl einer Funktionstyps

Allgemeine Vorgehensweise:

Um in einer Realsituation die math. Methode „Regression" sachgerecht angehen zu können sind folgende Schritte zu beachten (zwangsläufig auch für Schüler maßgebend!):

(1) Ist zur Mathematisierung der Realsituation die mathematische Methode „Regression" geeignet?

(2) Lassen die Rahmenbedingungen der Realsituation eine Vermutung über einen oder mehrere Funktionstypen zu.
z.B.: „Eingangs- und Ausgangsgrößen weitgehend linear", „Änderung nehmen zu oder nehmen ab", „Änderung ist proportional zum Ausgangswert"

Hinweis:

In den meisten „Anwendungsaufgaben" wird die Situation verkehrt:
Gegeben ist i.d.R. ein funktionaler Zusammenhang zwischen Eingangs und Ausgangswert (häufig in Form variablen darstellung → Funktionstyp!).
Der Schüler soll dann im Nachhinein „begründen", warum dieser Funktionstyp die Realsituation geeignet darstellt!
Häufig wird der Schüler dann nur der Realsituation die Eigenschaften der Funktion „zuschreiben"!

(3) Nach Durchführung der Regression(en) wird das/die Bestimmtheitsmaß/e R^2 (bzw. der/die Korrelationskoeffizient(en) K) verglichen bzw. absolut bewertet.
Größeres R^2 (bzw. K) gilt als „bessere" Mathematisierung.
$K \geq 0,8$ „gilt" als gute Annäherung

Selbst wenn die „Allgemeine Vorgehensweise" beachtet wird, bleibt ungewiss, ob nicht ein Funktionstyp existiert, der die Realsituation besser beschreibt
bzw., welche Eingangsgrößen für die Beschreibung wesentlich sind.

<u>Hinweise:</u>

(1) Bei gleicher Messwertgrundlage können grundlegend verschiedene Funktionstypen zu ähnlichen Bestimmtheitsmaßen führen!

(Folie 4)

(2) Ist die Zahl der Messungen n und die Zahl der Eingangsvariablen k gleichgroß, so wird die multiple lineare Regression zu einer Funktion führen, für welche die Unterschiede $\hat{u}_i = 0$ sind, für alle i (alle Messpunkte liegen auf dem Graphen von f); $R^2 = 1$.
Es handelt sich aber nur um eine „Scheinidentität", da weitere Messungen „weitab" von diesem Funktionsgraphen liegen können.
Darum: $n \gg k$

<u>Fragestellung:</u>

Gibt es eine <u>mathematisch begründete Methode</u>, welche im Falle multipler linearer Regression ein Entscheidungskriterium liefert, <u>welche</u> Zugangsvariablen <u>wesentlichen Anteil</u> an der Beschreibung der Realsituation haben?

<u>Vorgabe:</u>

n Messwertpunkte mit je k (Zugang-)Eingangsvariablen

<u>Ziel:</u>

Methode diejenigen Eingangsvariablen zu <u>eliminieren</u>, die <u>keinen „wesentlichen" Anteil an der Beschreibung</u> der Realsituation haben.

Da die n Messwertpunkte fest vorgegeben sind, wird sich die

Gesamtvariabilität $V = \sum_{i=1}^{n} (y_i - \bar{y})^2$ nicht ändern

Wohl aber die

Variabilität der Prognosewerte

$$\hat{V} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2,$$

weil je nach Funktionstyp sich die $\hat{y}_i$ ändern

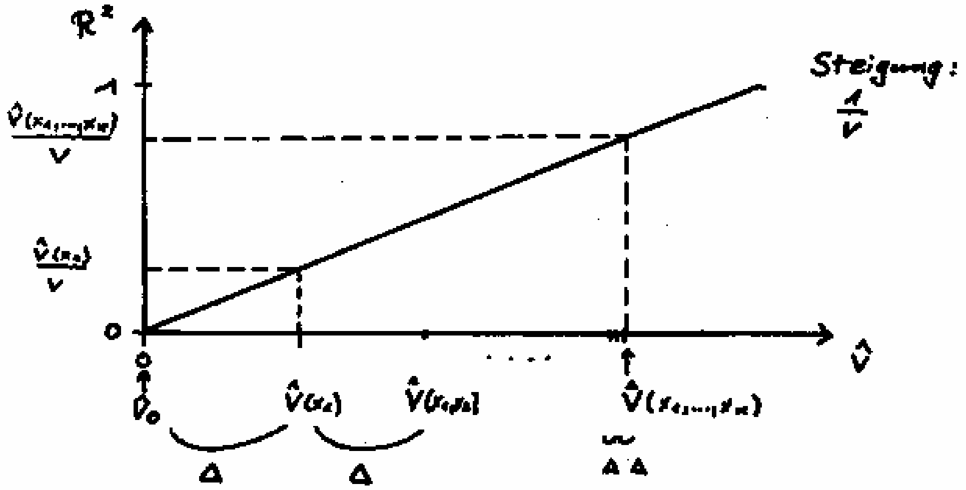
Damit ändert sich auch das

Bestimmtheitsmaß $R^2 = \frac{\hat{V}}{V}$

<u>Betrachtet man die $\hat{V}$ für die verschiedenen Funktionstypen:</u>

Funktionstyp	Regressionsfunktion		Variabilität der Prognosewerte	
$y = \beta_0$	$y = \hat{\beta}_0$	Zunahme $\downarrow$	$\hat{V} = \sum_i (\hat{\beta}_0 - \bar{y})$ $= \sum_i (\hat{\beta}_0 - \hat{\beta}_0)$ $= 0 =: \hat{V}_0$	<u>Hinweis:</u> Da $\bar{y} = \frac{1}{n} \sum y_i$ $= \frac{1}{n} \sum \hat{y}_i$ $= \frac{1}{n} \sum_i \hat{\beta}_0$ $= \hat{\beta}_0$
$y = \beta_0 + \beta_1 x_1$	$y = \hat{\beta}_0 + \hat{\beta}_1 x$		$\hat{V} := \hat{V}(x_1)$	
$\vdots$				
$y = \beta_0 + \beta_1 x_1 + \dots + \beta_K x_K$	$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_K x_K$		$\hat{V} := \hat{V}(x_1, \dots, x_K)$	

Graphische Darstellung von R^2 bei zunehmender Einbeziehen von Eingangsvariablen



An der graphischen Darstellung soll verdeütlicht werden, dass $x_1, x_2, \ldots$ nicht eliminiert woden dürfen wohl aber x_{k-1}, x_k

Allerdings:

Es könnte durchaüs sein, dass – nachdem x_1, x_2 bereits einbezogen würden – x_3 nür geringen „Erklärungsanteil" liefert und damit eliminiert wird.

Bei gleicher Situation aber – nachdem x_2 und x_3 bereits einbezogen wurden – x_1 nür geringen Erklärungsanteil liefert.

Eine Methode, die alle diese Kombinationen systematisch berücksichtigt ist unter Oberguggenberger, Ostermann, Unterweger; Lineare Regression, 2004, S. 12 (Internetstichwort: „Universität Innsbruck Lineare Regression") nachzulesen

– 35 –

VII Regression bei „nichtlinearen" Funktionstypen

Es wird davon ausgegangen, dass der betrachtenswerte Funktionstyp <u>nicht</u> dem Typ einer (multiplen) linearen Funktion entspricht; als <u>nicht</u> von der Form

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

oder allgemeiner

$$T_0(y) = \beta_0 + \beta_1 T_1(x_1) + \dots + \beta_k T_k(x_k) \quad \text{ist.}$$

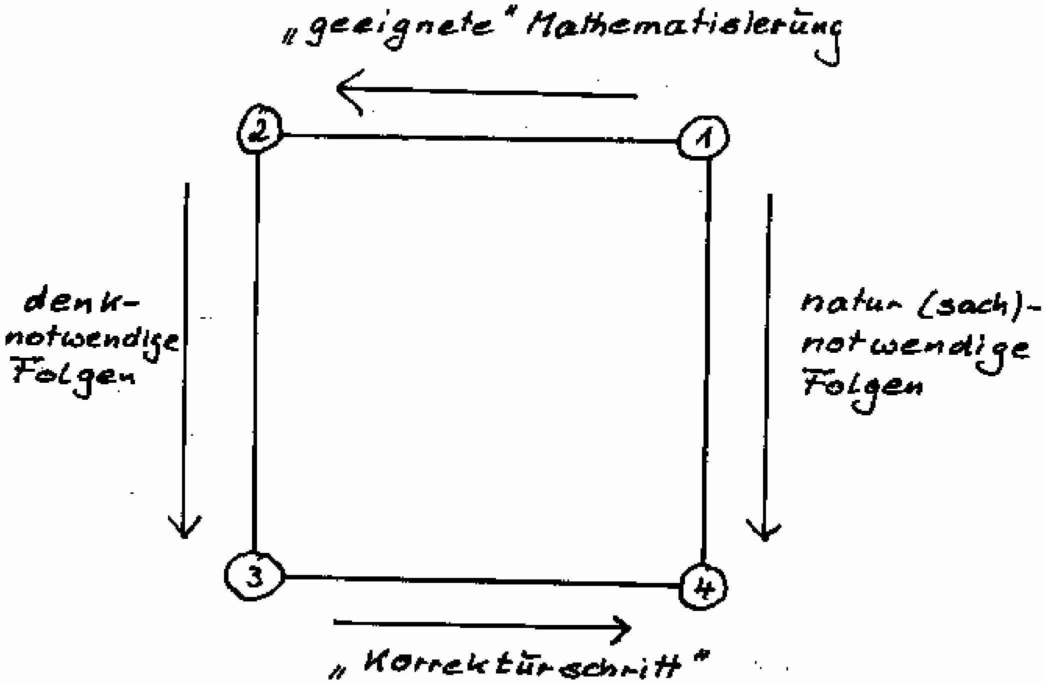
Insbesondere der <u>logistische Funktionstyp</u> lässt sich nicht linear darstellen.

Eine ausführliche Beschreibung <u>nichtlinearer Regressionen</u> (z.B. logistisch) durch <u>Iteration</u> wird in
Universitäts-Rechenzentrum Trier; Nichtlineare Regression mit SPSS, 1998 (Internetstichwort: „Universität Trier Nichtlineare Regression")
beschrieben

(Hier wird wohl die Regression empirisch interpretiert; das Verfahren selbst ist aber auch auf deterministische Betrachtungsweisen anwendbar)

(Folie 1)

Helmholtz - Diagramm



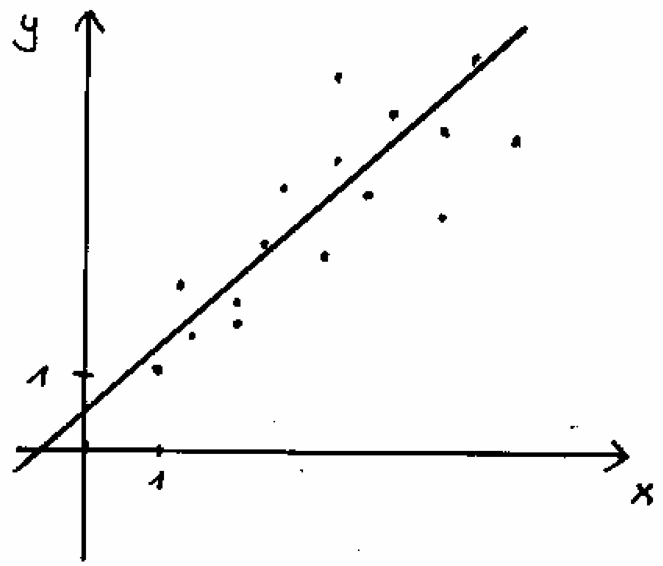
(Folie 2)

Grundgedanke „Regression"

Zahlenpaare (Messungen)
(1; 2,5), (1,7; 3,6), ... (x;y) $\Longrightarrow$
gegeben

f: Funktionaler Zusammenhang mit
$y \approx f(x)$ für alle Paare
gesucht

graphische Darstellung:



(Folie 3)

Realsituation	**Mathematisierte Form**
1. Messwertpaare sind vorgegeben; gesucht ist ein funktionaler Zusammenhang, der jedem x-Wert **genau** seinen y-Wert zuordnet. Ein Verfahren, das insofern ökonomisch sein kann, weil an die Stelle der zu „merkenden“ y -Werte nur das formale Verfahren tritt mit dessen Hilfe die y-Werte bei Bedarf berechnet werden können.	Polynom oder andere Funktion, deren Koeffizienten eindeutig durch die Wertepaare bestimmt werden. **Regression ungeeignet**, da y-Werte und f(x)-Werte nur angenähert übereinstimmen.
2. Von allen x-Werten soll **bekannt** sein, welcher **funktionale Zusammenhang** verfahrensmäßig zu den y-Werten besteht; zudem wird davon ausgegangen, dass aufgrund der Situation bekannt ist, welchem **Standard-Typ** diese Funktion **ähnelt** (Polynom, trigonometrische Funktion, etc.); gesucht ist eine **eindeutige** Modifikation des Standard-Typs, welcher den vorgegebenen funktionalen Zusammenhang beliebig genau darstellen kann.	Taylor-Polynom, harmonische Analyse Für eine **Regression** nicht uninteressant, da ja nur ungefähre Übereinstimmung von y-Werten und f(x)-werten gefordert wird; aber **ungeeignet**, weil bei der Regression der funktionale Zusammenhang gerade nicht vorgegeben, sondern gesucht wird.
3. Von einem System soll **kein funktionaler Zusammenhang** aber **gewisse Symmetrieeigenschaften** bekannt sein („die Augenzahlen eines Würfels kommen gleichwahrscheinlich vor“); das System soll im Hinblick auf Versuche (x-Werte) und deren Ergebnisse (y-Werte) funktional beschrieben werden.	Methoden der Statistik. Für die **Regression** zunächst nicht uninteressant, weil ja funktionale Größen wie der Erwartungswert damit beschrieben werden könnte aber insofern **ungeeignet** weil von der Regressionsfunktion ja gerade erwartet wird, dass im **Wiederholungsfall** bei gleichem x-Wert der y-Wert sich möglichst wenig vom f(x)-Wert unterscheidet.

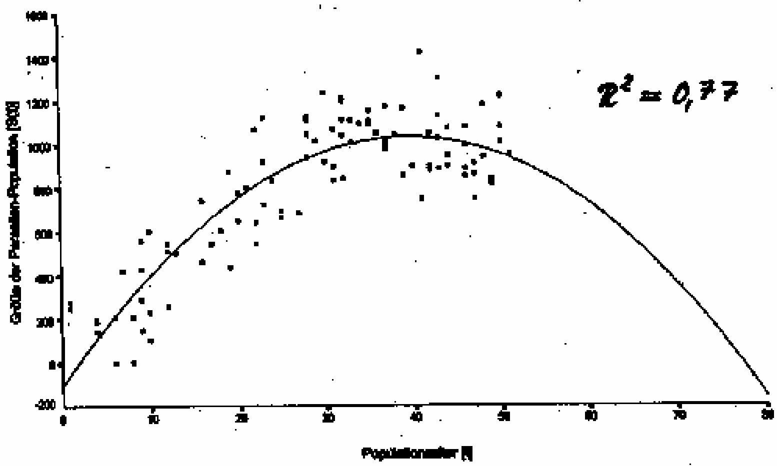
Realsituation	***Mathematisierte Form***
4. Von einem System sollen ihre **Teilsystemen** (funktional) **bekannt** sein, aus denen sich das Gesamtsystem zusammensetzt (z.B. eines Hubsystems in ein Drehsystem wie im Kolbenmotor); gesucht ist ein funktionaler Zusammenhang zwischen Eingangsgrößen (x-Werte) und Ausgangsgrößen (y-Werte).	Koppelung der Funktionen der Teilsysteme durch Gleichungen etc. . Für die **Regression** ungeeignet, da von den (x,y)-Wertepaaren nur die x-Werte als kontinuierliche Größe angebbar sind und der funktionale Zusammenhang sich **eindeutig** aus den Teilsystemen ergeben kann.
5. Von einem System sollen einzelne **bestimmte Wertepaare** bekannt sein und bestimmte **„lokale"Eigenschaften** (z.B. „Änderungsrate); gesucht ist ein funktionaler Zusammenhang, der die genannten Bedingungen berücksichtigt.	Differentialgleichungen Für die **Regression** nicht uninteressant, weil bestimmte Wertepaare vorgegeben sind; aber **ungeeignet**, weil bestimmte Strukturen des Systems aber- mit wenigen Ausnahmen - keine y-Werte bekannt sind. Diese „Strukturvorgabe" ist allerdings bei der Regression noch mal aufzugreifen, weil in der allgemeinen Regression die Frage aufzugreifen ist, **durch welcher Funktionstyp die y-Werte anzunähern sind!**
5. Von dem System sind (x,y)-Werte bekannt aber das **System selbst ändert** sich (in bekannter oder unbekannter Weise) bei jeder Messung; gesucht sind Zusammenhänge.	Sog. mehrstufige Prozesse, chaostheoretische Betrachtungen; suche nach Stabilitätspunkten. Für **Regression** ungeeignet, da an die Regression ja gerade die - bisher ungenannte Forderung - gestellt wird dass bei wiederholter Vorgabe bestimmter x-Werte die y-Werte durch die Regressionsfunktion zumindest näherungsweise vorausgesagt wird.

Realsituation	*Mathematisierte Form*
6. Bekannte Zusammenhänge zwischen **irgendwelchen** Objekten sollen systematisiert werden.	Aus Relationen („Ähnlichkeitsrelationen“) werden über eine Analyse Klassen äuquivalenter gebildet. Für die **Regression** ungeeignet, da für die x, y-Werte Größen vorausgesetzt werden, um beispielsweise die Forderung nach einer „möglichst geringen Unterscheidung“ zwischen y - und f(x)-Werten mathematisch fassen zu können.

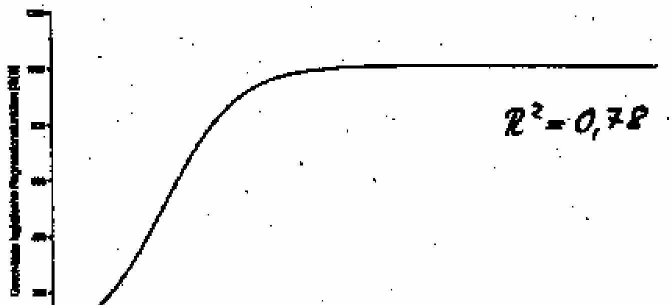
(Folie 4)

Größe einer Parasitenpopulation auf einem Wirtstier in Abhängigkeit vom Populationsalter

Funktionstyp: Quadratische Funktion



Funktionstyp: Logistisches Wachstum



Aus:
Universitäts-Rechenzentrum Trier: NichtLineare Regression mit SPSS, 1998 (Internetstichwort: „Universität Trier Nichtlineare Regression"), S. 6, 10