

Kurvenanpassung in EXCEL

Inhalt

1	Anpassungsverfahren.....	3
1.1	Ausgleichsgerade (Lineare Regression)	3
1.1.1	Verfahren	3
1.1.2	Korrelationskoeffizient R und Bestimmtheitsmaß R^2	4
1.1.3	Tabellenkalkulation	5
1.2	Unterscheidung der weiteren Verfahren	6
1.3	Glättung.....	7
1.3.1	Grundlagen	7
1.3.2	Gleitender Mittelwert	8
1.3.3	Exponentielle Glättung	8
1.3.4	Median-Glättung	8
1.3.5	Savitzky-Golay-Glättung.....	8
1.4	Averaging	10
1.5	Interpolation	11
1.5.1	Grundlagen	11
1.5.2	Lineare Interpolation	11
1.5.3	Polynom-Interpolation.....	13
1.5.4	Spline-Interpolation.....	13
1.6	Approximation	14
1.6.1	Grundlagen	14
1.6.2	Direkte Verfahren	14
1.6.3	Linearisierung	15
1.6.4	Iterative Verfahren.....	15
2	Neuabtastung	16
3	Kurvenanpassung mit Standardfunktionen in EXCEL	17
3.1	Lineare Funktion.....	17
3.2	Polynomfunktion.....	17
3.3	Potenzfunktion	18
3.4	Exponentialfunktion.....	19
3.5	Logarithmusfunktion	20
3.6	Gleitender Durchschnitt.....	21

Kurvenanpassung in EXCEL

4	Kurvenanpassung mit dem SOLVER.....	22
4.1	Iterative Verfahren (SOLVER).....	22
4.2	Durchführung der Berechnung.....	22
4.3	Beispiel aus der Spektroskopie	23
4.3.1	Ansatz	23
4.3.2	Realisierung der Anpassung.....	23
4.4	Weiteres SOLVER-Beispiel: Potentiometrische Titration	26
5	Übungsblätter	27
5.1	Lineare Kalibrierfunktion	27
6	Hinweise	28
6.1	Dokumentenstruktur	28
6.2	Text-Dateien	28
6.3	EXCEL-Dateien	28

Anpassungsverfahren

1.1 Ausgleichsgerade (Lineare Regression)

1.1.1 Verfahren

Die Ausgleichsgerade ist ein Ausgleichs-Verfahren zur Kurvenanpassung (Approximation). Durch ein Feld von Datenpunkten in einem Diagramm wird eine Gerade gelegt, die den Trend (steigend, fallend) anzeigt. Diese Gerade wird **Ausgleichsgerade** oder auch **Regressionsgerade** oder **Lineare Regression** genannt. Die Gerade lässt sich mathematisch durch die Funktionsgleichung $y = b \cdot x + a$ beschreiben. b ist die Steigung, a der y-Abschnitt.

Achtung: In der Schule wird meist m für die Steigung und b für den y-Achsenabschnitt verwendet. In technischen Dokumenten und Normen findet man meist die oben angegebenen Bezeichnungen b für die Steigung und a für den y-Achsenabschnitt.

Die Ausgleichsgerade wird so durch ein Punktfeld gelegt, dass sie allen Punkten "möglichst nahe" kommt. Zur Bestimmung der Lage einer Ausgleichsgeraden könnte man einen mechanischen Analogrechner bauen:

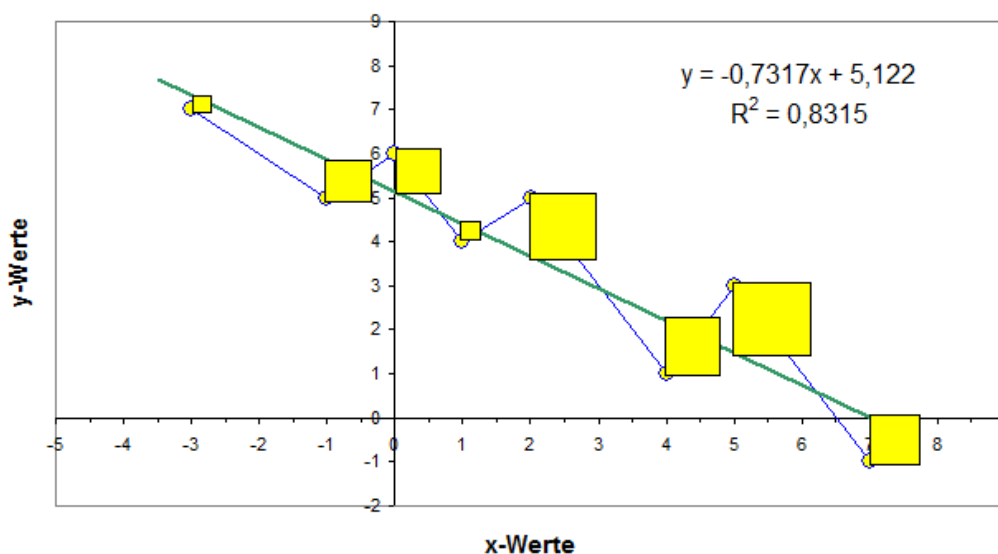
Mechanischer Analogrechner: Auf eine Holzplatte wird das Koordinatensystem gezeichnet und an die Stellen der Datenpunkte Nägel eingeschlagen. Ein starrer Stab wird mit Gummiringen mit den Nägeln verbunden. Der Stab stellt sich so ein, dass die Spann-Energie der Gummiringe minimal wird. Die Spann-Energie ist proportional zum Quadrat des Abstandes. Der Stab zeigt dann die Ausgleichsgerade an.

Das Problem der Minimierung der Summe der y-Abstandsquadrate lässt sich mit Hilfe der Differentialrechnung lösen und führt auf geschlossene (direkt berechenbare) Formeln für Steigung b und Abschnitt a . Die Summen gehen jeweils über alle Koordinatenwerte.

$$b = \frac{\sum_{i=1}^n xy - (\sum x)(\sum y)}{\sum_{i=1}^n x^2 - (\sum x)^2}$$

mit \bar{X} : Mittelwert aller x -Werte und \bar{Y} : Mittelwert aller y -Werte erfolgt die Berechnung des Achsenabschnitts a (mit der bereits berechneten Steigung b) über $a = \bar{Y} - b \cdot \bar{X}$.

Der Schnittpunkt der Geraden mit der x -Achse lässt sich aus $x_0 = -a/b$ berechnen (a : y-Abschnitt, b : Steigung). Die Mittelwerte der x -Daten \bar{x}_s bzw. der y -Daten \bar{y}_s ergeben die Koordinaten des Schwerpunkts (\bar{x}_s/\bar{y}_s) , der auf der Ausgleichsgeraden liegt.



In manchen Fällen wird eine Anpassungsgerade gesucht, die auf jeden Fall durch den Nullpunkt des Koordinatensystems geht. Das Rechenverfahren muß dann also nur die Steigung bestimmen.

Kurvenanpassung in EXCEL

Achtung: Vertauscht man die x- und y-Werte und berechnet davon die Ausgleichsgerade, erhält man eine andere Gerade als wenn man die erste Geradengleichung nach x auflöst und dann x und y vertauscht.

Man unterscheidet also die "Regression von y auf x" und "Regression von x auf y". Der Unterschied der dabei berechneten Geraden wird durch die Lineare Korrelation ausgedrückt.

1.1.2 Korrelationskoeffizient R und Bestimmtheitsmaß R^2

Um eine Ausgleichsgerade können sich die ursprünglichen Daten eng gruppieren oder auch weit streuen. Der Geradengleichung lässt sich nicht ansehen, wie eng Daten und Gerade miteinander zusammenhängen (korrelieren).

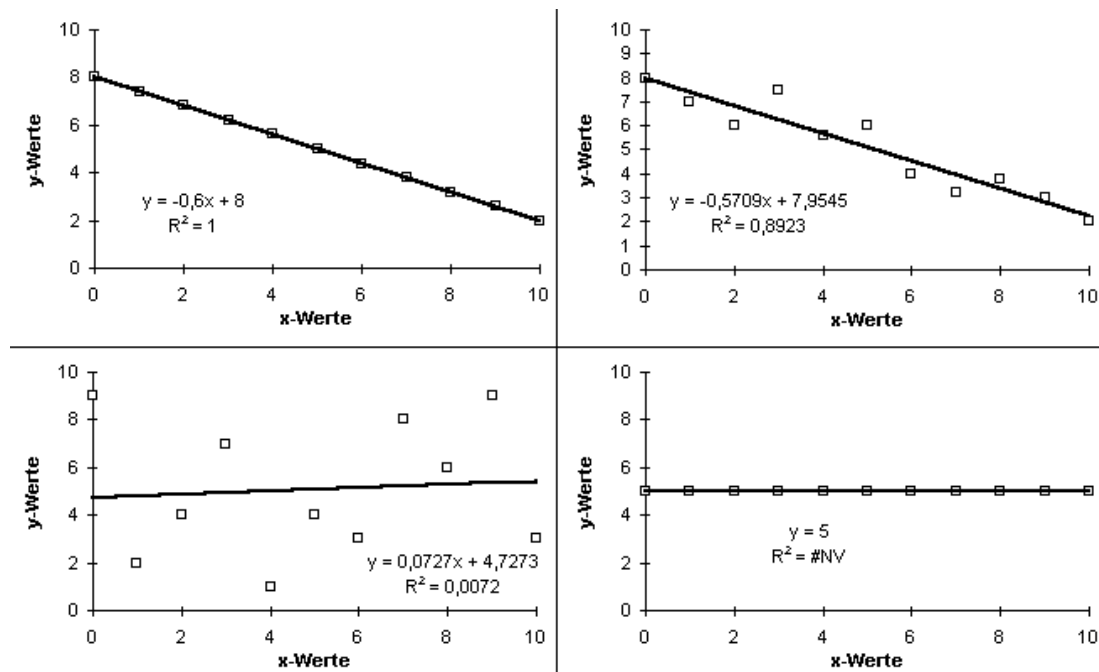
Die Eigenschaft, dass sich beim Vertauschen der x- und y-Werte verschiedene Geraden ergeben, kann man benutzen, um eine Größe zu berechnen, die den Zusammenhang, die Korrelation, misst. Beide Geraden gehen durch den Schwerpunkt und bilden eine Schere, die umso mehr geschlossen ist, je enger sich die Datenpunkte um jede einzelne Gerade gruppieren. Liegen alle Punkte schon auf einer Geraden, dann fallen beide Geraden zusammen, die Schere ist geschlossen.

Der berechnete Wert, der (Pearsonsche) Korrelationskoeffizient R, liegt immer zwischen -1 und +1. Bei $R = 1$ liegen alle Punkte auf der Geraden und die Steigung ist positiv. Bei $R = -1$ liegen alle Punkte auf einer Geraden mit negativer Steigung. Bei $R = 0$ ist entweder die Steigung $b=0$ (in EXCEL #NV: not valid) oder die Punkte streuen so in der x-y-Ebene, dass keine Vorzugsrichtung erkennbar ist. Zwischen diesen Fällen sind alle Zwischenwerte möglich.

Das Bestimmtheitsmaß (oder Determinationskoeffizient) R^2 ist das Quadrat des Korrelationskoeffizienten R. Es liegt immer zwischen 0 und 1 und enthält nicht mehr die Information über das Vorzeichen der Steigung.

$$-1 \leq R \leq +1 \quad 0 \leq R^2 \leq 1 \quad R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2] * [n \sum y^2 - (\sum y)^2]}}$$

es wird jeweils über alle Indices summiert.



Erstaunlich ist, dass im vierten Beispiel kein Ergebnis angegeben werden kann. Mathematisch ist dieser Fall definiert mit $R=R^2=0$, aber im Programmcode einiger Tabellenkalkulationen nicht

Kurvenanpassung in EXCEL

umgesetzt. $R=R^2=0$ besagt, dass von keiner Abhängigkeit der y-Werte von den x-Werten ausgegangen werden kann.

Für andere Fragestellungen gibt es weitere Korrelationsmaße (z.B. Spearman, Kendall).

1.1.3 Tabellenkalkulation

In der Tabellenkalkulation mit den x-Werten in den Zellen A6 bis A12 und den y-Werten in den Zellen B6 bis B12, sehen die zugehörigen Funktionen so aus:

=STEIGUNG(B6:B12;A6:A12)

=ACHSENABSCHNITT(B6:B12;A6:A12) (jeweils zuerst die y-Werte, dann die x-Werte)

Bestimmtheitsmaß und Korrelationskoeffizient werden durch folgende Funktionen berechnet:

=BESTIMMTHEITSMAS(B6:B12;A6:A12)

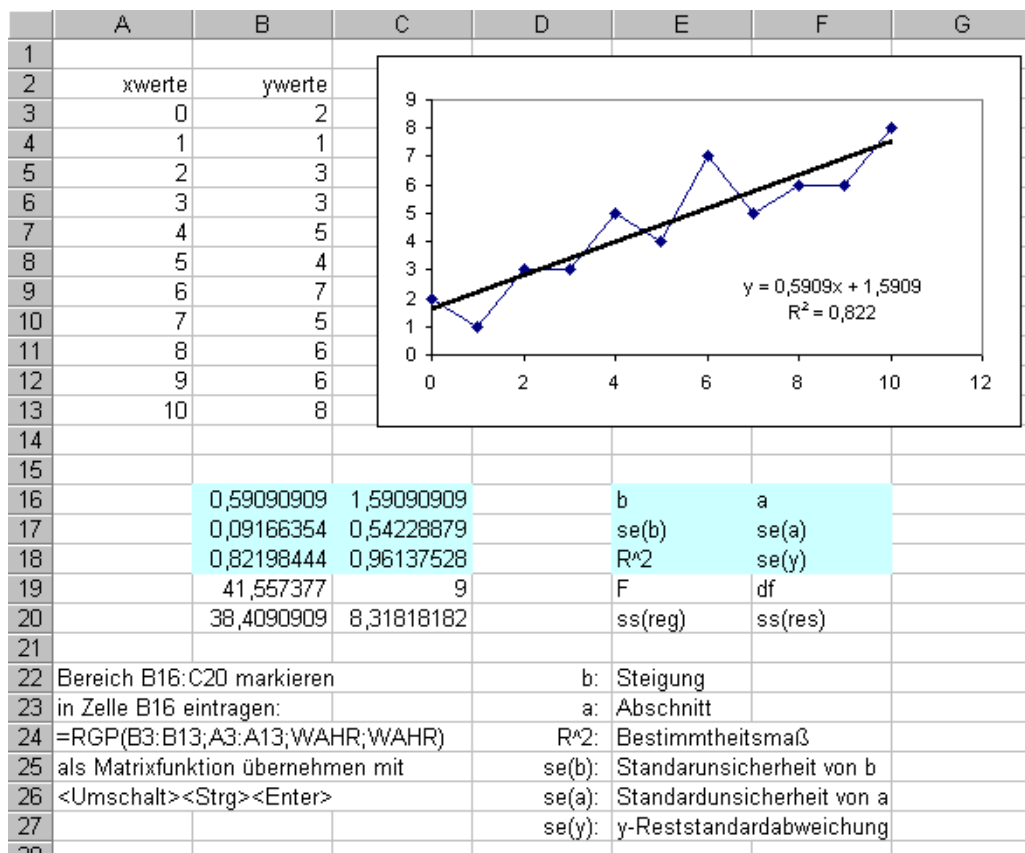
=KORREL(B6:B12;A6:A12)

Weitere Berechnungen

Die Abweichungen von der Geraden werden als zufällige (normalverteilte) Schwankungen interpretiert. Damit unterliegen auch die Steigung b und der Abschnitt a gewissen Unsicherheiten. Die Statistik erlaubt dann die Berechnung der Standardunsicherheiten Δb für b und Δa für a. In der Tabellenkalkulation erhält man die Werte z.B. mit der Statistik-Funktion RGP.

=RGP(ywerte;xwerte;WAHR,WAHR)

Für Spezialisten: $se(y)$ ist die Reststandardabweichung, also die Standardabweichung aus der Streuung der Punkte um die Ausgleichsgerade. Die Größen $se(b)$ und $se(a)$ sind abhängig von der Anzahl der Datenpunkte und werden mit Hilfe der t-Verteilung berechnet.



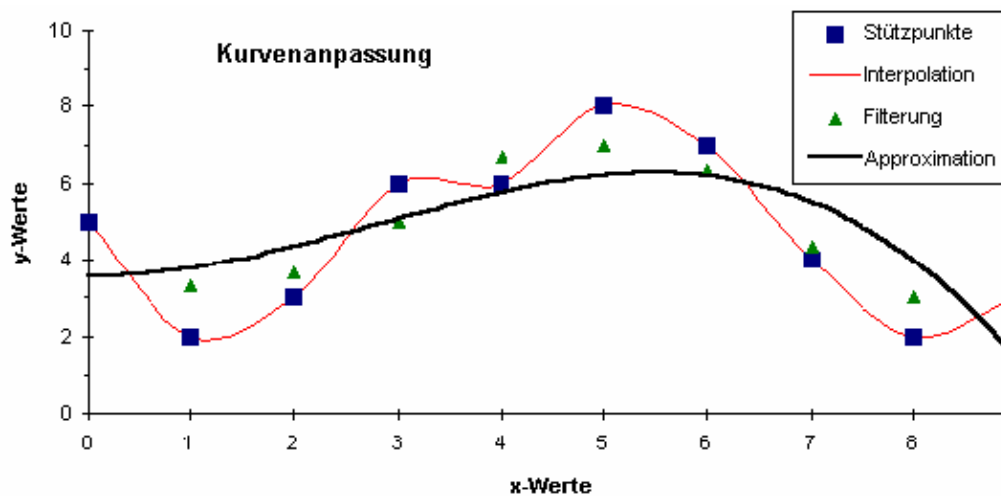
1.2 Unterscheidung der weiteren Verfahren

Ergebnisse von Messungen werden in der Form von numerischen Daten gespeichert. Alle Informationen, die man über das Messobjekt und den Messprozess gewinnen kann, sind damit in Zahlenfolgen enthalten.

Kontinuierliche Zusammenhänge zwischen Messgrößen können nur zeit- und wertdiskret erfasst werden. Grafische Darstellungen führen so zu "Punktwolken". Die Auswahl eines geeigneten Verfahrens zur Beschreibung der Punktwolke durch eine Kurve hängen von den Vorkenntnissen über den beobachteten Prozess, vom Messvorgang und dem Verwendungszweck der Anpassung ab. Kurvenanpassungen stellen daher immer auch eine Interpretation der Daten dar.

Bei der **Glättung** (ein Spezialfall der **Filterung**) werden keine (stetigen) Funktionen oder Funktionsabschnitte berechnet, sondern nur Ersatzwerte für die Stützpunkte.

Beim **Averaging** wird ein wiederholbarer Vorgang mehrfach ausgeführt. Die Abtaststellen müssen bezüglich des Vorgangs immer gleich sein. An allen Abtaststellen wird der Mittelwert (evtl. auch der Median) über alle Wiederholungen berechnet. Zufällige Schwankungen (Rauschen) werden dadurch reduziert.



Bei der **Interpolation** werden Funktionswerte für das Intervall zwischen je zwei bereits vorhandenen Stützpunkten der Kurve berechnet. Im Allgemeinen müssen dabei für jedes Intervall die Parameter (Konstanten) der Interpolationsfunktion neu berechnet werden.

Bei der Auswahl einer geeigneten Interpolationsfunktion sind neben dem Funktionstyp wie Gerade, Polynom oder Exponentialfunktion auch die Anschlussbedingungen (z.B. kein Sprung oder Knick) und die Randbedingungen (z.B. Steigung und Krümmung am Kurvenanfang und -ende) zu beachten.

Die Begriffe **Approximation**, **Ausgleich**, **Anpassung**, **Regression** und **Fit** werden häufig nicht unterschieden. Das Ergebnis ist dabei jeweils eine Kurve, von der im Gegensatz zur Interpolation nicht mehr verlangt ist, dass sie die Stützpunkte enthält. Statt dessen werden durch die Ausgleichsfunktion Forderungen erfüllt wie "minimale Gesamtabweichung bei gegebenem Funktionstyp" oder "maximale Abweichung muss unterhalb einer vorgegebenen Schranke liegen". Man kann sich vorstellen, dass die Kurve von den Datenpunkten angezogen wird. Als Ausgleichsfunktion kommen die gleichen Funktionstypen wie bei der Interpolation in Betracht.

Mit der Bezeichnung "Fit" bezieht man sich oft auf ein iteratives Verfahren. In der Mathematik wird Approximation oft als Annäherung an eine bekannte Funktion verstanden, was in der Messtechnik jedoch nur selten vorkommt.

Bei geringer Anzahl von Stützpunkten geht die Approximation in die Interpolation über (wenn die Anzahl der Stützstellen gleich der Anzahl der Funktionsparameter ist).

Bei einer **Neuabtastung** wird zuerst eine Interpolation oder Approximation berechnet. Die Interpolations- bzw. Approximationsfunktion wird anschließend an festgelegten neuen Stützpunkten abgetastet. Ziel ist eine äquidistante Abtastung, gleicher Taktabstand oder gleiche Datenanzahl bei verschiedenen Signalen.

1.3 Glättung

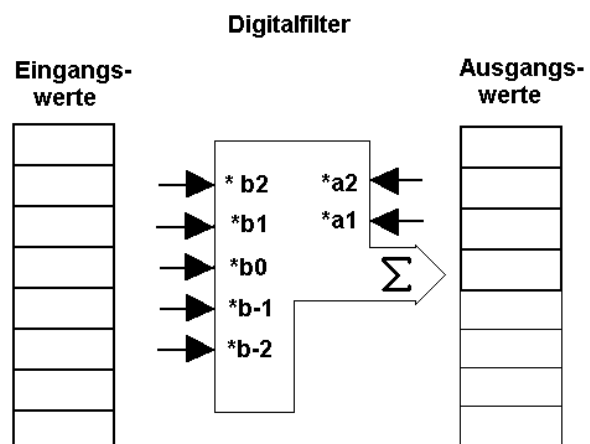
1.3.1 Grundlagen

In der Messtechnik werden Datenreihen erfasst, die entweder die zeitliche Folge von Messwerten einer physikalischen Größe (z.B. Änderung des Drucks von der Zeit bei einem Pressvorgang) oder die Abhängigkeit zweier oder mehrerer physikalischer Größen (z.B. Abhängigkeit der Viskosität von der Temperatur) darstellen. Die Messdaten sind häufig durch zusätzliche Einflüsse verfälscht. Diese lassen sich unterscheiden in systematische Einflüsse (z.B. Schwankung der Helligkeit einer Lampe im Rhythmus der Netz-Wechselspannung) und in zufällige Einflüsse (z.B. Erschütterungen, Wärmebewegung der Atome). Kennt man charakteristische Eigenschaften dieser störenden Einflüsse, kann man sie durch Filter eliminieren oder zumindest abschwächen.

Zufällige Einflüsse bezeichnet man auch als Rauschen. Rauschen ist in den meisten Messsignalen enthalten. Die Verringerung des Rauschens bezeichnet man als Glättung. In diesem Kapitel werden spezielle Glättungsverfahren behandelt. Die nächsten beiden Abschnitte beschreiben jedoch Sachverhalte, die über den Spezialfall der Glättung hinaus gehen.

Arbeitsweise eines Digitalfilters

Das zu filternde Signal liegt in Form einer Liste mit numerischen Werten vor. In der Regel ist davon auszugehen, dass die Daten äquidistant (z.B. in gleichen Taktintervallen) abgetastet wurden. Bei der Filterung wird eine neue Datenliste erzeugt, bei der jeder einzelne Wert aus einer Gruppe von Werten der Originaldaten berechnet wurde. Ein Filter wird durch die Rechenvorschrift bestimmt, nach der die neuen Werte berechnet werden. Bei sog. Linearen Filtertypen werden die Berechnungen durch eine Lineare Funktion (Geradengleichung, Addition und Multiplikation) durchgeführt. Hier genügt die Angabe einer Liste von Filterkoeffizienten (Kern, kernel). In Berechnungsformeln nichtlinearer Filter kommen auch nichtlineare Funktionen (z.B. Polynome, Absolutbetrag, Median) vor. Die Bestimmung der erforderlichen Filterungsregeln bzw. der erforderlichen Koeffizienten ist meist sehr rechenaufwändig. Die Anwendung auf einen Datensatz erfordert oft nur die Grundrechenarten.



Glättung als Spezialfall

Bei der Erfassung einer Messdatenreihe treten häufig zufällige Schwankungen auf. Glättungsverfahren haben den Zweck, diese zufälligen Anteile in der Messreihe zu reduzieren. Glättungsverfahren sind daher Spezialfälle der Filterung. Für jeden einzelnen Messwert muss dazu eine Rechnung ausgeführt werden. Der zu bearbeitende Datenpunkt wird mit einigen seiner Nachbarwerte verglichen und daraus der korrigierte Datenpunkt berechnet. Glättung ist eine digitale Tiefpass-Filterung.

Bei der Glättung werden keine Datenpunkte zwischen den Abtastwerten berechnet, im Gegensatz zu Interpolation und Ausgleich. Da die Glättungsverfahren am Anfang und am Ende der Datenreihe oft nicht definiert sind, werden dort dann keine Glättungswerte berechnet. Dann muss darauf geachtet werden, dass die Zuordnung der Werte (x-Werte zu Original und Glättung) erhalten bleibt.

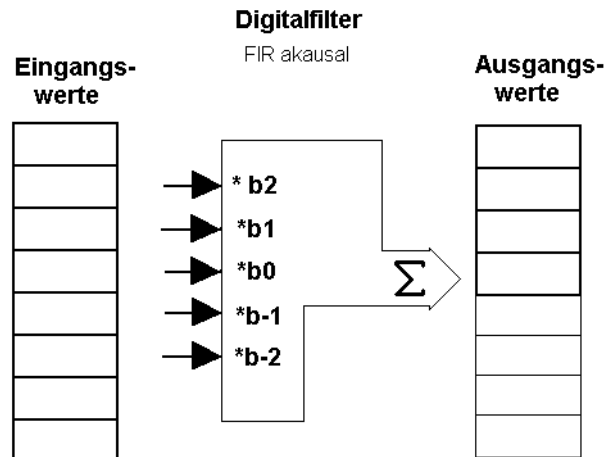
Die folgenden Verfahren gehen von einer äquidistanten Abtastung aus (gleiche Abstände in Zeit- bzw. x-Richtung). Sie lassen sich aber in ähnlicher Form auch auf andere Datenreihen anwenden. Ein mehrmaliges Anwenden eines Filters erhöht die Glättungswirkung. Sie kann aber auch auf einmal durch Verwendung eines entsprechenden Filters höherer Ordnung erreicht werden.

1.3.2 Gleitender Mittelwert

Berechnet man für jeden Messwert den Mittelwert, der sich aus dem Messwert und seinen beiden Nachbarwerten ergibt, erhält man das am häufigsten eingesetzte Glättungsverfahren, ein einfaches FIR-Filter. Man geht davon aus, dass die Abweichungen nach oben und unten zufällig verteilt sind, und sich dadurch in den meisten Fällen reduzieren.

Um einen gefilterten Wert zu berechnen, werden die mit den Pfeilen versehenen Eingangswerte mit den Filterkoeffizienten b multipliziert und addiert.

Im einfachsten Fall haben hier alle Koeffizienten den Wert 0,2. Dann ergibt sich der arithmetische Mittelwert aus den hier fünf Eingangswerten.



Statt dreier benachbarter Datenwerte können für die Mittelwertbildung auch größere Umgebungen verwendet werden. Typisch sind 3, 5, 7 oder 9 Punkte. Das Ergebnis wird immer "glatter" und kann damit auch wichtige Details "verschmieren". Durch die Wahl unterschiedlicher Filterkoeffizienten werden die Glättungseigenschaften weiter beeinflusst.

Für das Verfahren werden auch die Bezeichnungen verwendet: gleitender Durchschnitt, moving average, FIR-Filter (Finite Impuls Response), Transversalfilter, nichtrekursives Filter.

1.3.3 Exponentielle Glättung

Bei der exponentiellen Glättung wird im einfachsten Fall der letzte geglättete Wert und der neue Messwert gemittelt. Dies führt bei einem pulsformigen Eingangssignal auf ein exponentiell abfallendes Ausgangssignal. Durch eine Gewichtung der Rechenwerte lässt sich die Abfall-Geschwindigkeit beeinflussen (z.B. $y_i = 0,3 \cdot x_i + 0,7 \cdot y_{i-1}$).

Für das Verfahren werden auch die Bezeichnungen verwendet: rekursives Filter, IIR-Filter (Infinite Impuls Response), Autoregressives Filter.

1.3.4 Median-Glättung

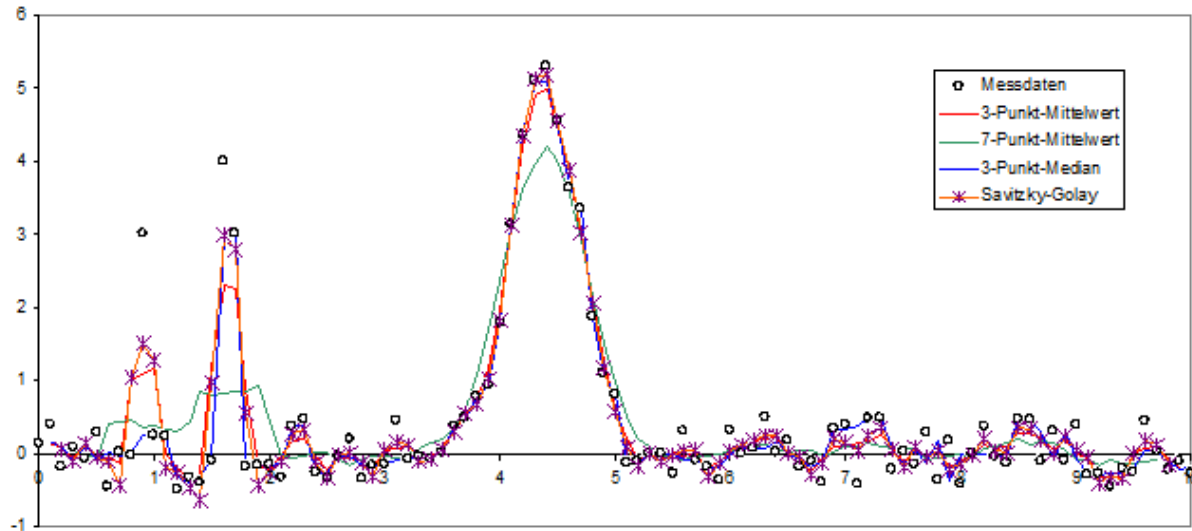
Zur Elimination von "Ausreißern" wird jeder Wert statt durch den Mittelwert durch den Median seiner Umgebung ersetzt. Der Median einer Datengruppe ist der Wert, von dem aus gesehen höchstens die Hälfte der Werte größer und höchstens die Hälfte der Werte der Datengruppe kleiner ist als der Medianwert (z.B. $\text{median}(7, 3, 4) = 4$).

Bildet man den Median aus jeweils drei benachbarten Werten, dann verschwinden einzelne Ausreißer, wie sie bei seltenen, zufälligen Schwankungen meist auftreten. Ausreißerpaare in gleicher Richtung bleiben erhalten. Die Median-Glättung (oder Spike Elimination) wird auch in der Bildbearbeitung eingesetzt, um einzelne helle oder dunkle Punkte zu eliminieren.

EXCEL-Blatt dazu: median.xls

1.3.5 Savitzky-Golay-Glättung

Bei der Auswertung von Spektren mit verrauschten Spektrallinien unterschiedlicher Breite führt die Tiefpass-Filterung dazu, dass sich die Linien verbreitern und in der Höhe reduziert werden. Bei der Savitzky-Golay-Glättung werden die Filterkoeffizienten eines FIR-Filters so bestimmt, dass ein Ausgleichspolynom bestimmt und damit der gefilterte Wert berechnet wird. Das Verfahren ermöglicht auch die Berechnung von Zwischenwerten, so dass bei sehr schmalen Peaks das Maximum der Glättung über den Datenpunkten liegen kann.



Das Savitzky-Golay-Glättungsverfahren wird auch "DISPO: Digital Smoothing Polynomial Filter" genannt.

Wie bei der Mittelwertglättung wird die Berechnung mit speziellen Koeffizienten durchgeführt. In der Literatur findet man für typische Filterlängen und Polynomgrade die zugehörigen Filterkoeffizienten.

Beispiel: Filterkoeffizienten für Ausgleichspolynome 2. Grades beim Filtern über 5 Punkte:

-0.086 0.343 0.486 0.343 -0.086

Ein Datengenerator für einen verrauschten Peak zur Untersuchung von Glättungsverfahren steht als Tabellenblatt unter glättbeispiele.xls.

Ergänzung: Aufgaben eines Filters

Bei der Filterung werden bestimmte Signaleigenschaften verstärkt oder abgeschwächt. Meist wird die Zusammensetzung eines Signals bezüglich seiner Frequenzanteile verändert. Mit einem Filter lassen sich bestimmte Frequenzanteile verstärken und andere abschwächen. Jedes Signal kann man sich zusammengesetzt denken aus Sinus- und Cosinusschwingungen verschiedener Frequenzen. Daher spielt der Wechsel zwischen Zeit- und Frequenzbereich (Fourier-Transformation) bei der Beschreibung von Filtern eine große Rolle.

Typische Filterfunktionen aus der Signaltechnik sind:

- Hochpass (Unterdrückung tiefer Frequenzen, hohe Frequenzen "passieren" das Filter)
- Tiefpass (Unterdrückung hoher Frequenzen, tiefe Frequenzen "passieren" das Filter)
- Bandpass (Durchlass eines zweiseitig begrenzten Frequenzbereichs)
- Bandsperre (Unterdrückung eines zweiseitig begrenzten Frequenzbereichs)
- Allpass (Beeinflussung der Phasen)

Bei der Messdatenverarbeitung sind zusätzlich die folgenden Verfahren mit speziellen Filtereigenschaften gebräuchlich:

- Integration
- Differentiation
- harmonisches Oversampling
- Modulation und Demodulation
- Frequenzverschiebung
- Hüllkurve

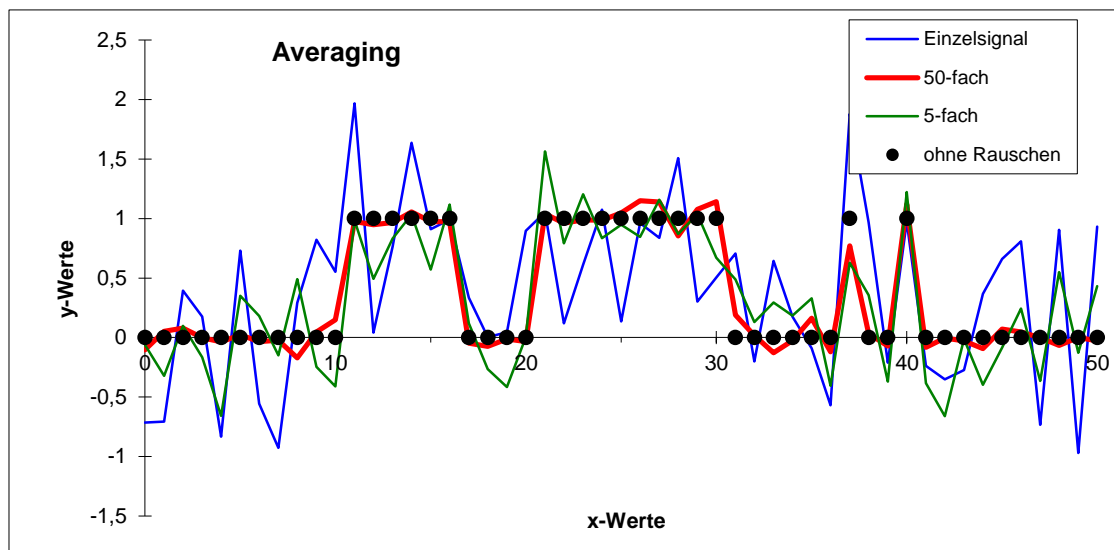
1.4 Averaging

Beim Averaging werden durch die mehrfache Wiederholung der Messung zufällige Schwankungen ausgeglichen. Während der eigentliche Signalanteil immer gleich bleibt, schwankt der zufällige Anteil in Größe und Richtung.

Addiert man alle Signale, dann addieren sich die eigentlichen Signalanteile, während sich die zufälligen Anteile teilweise wieder aufheben.

Dividiert man die Summe der n Signale wieder durch n, ist der störende, zufällige Anteil reduziert.

Nach den Gesetzen der Statistik wird dieser Anteil um den Faktor \sqrt{n} reduziert. Bei z.B. 25-maliger Wiederholung der Messung erhält man eine Reduktion auf 1/5.



In der EXCEL-Datei „average.xlsx“ wird das Averaging-Verhalten dargestellt.

1.5 Interpolation

1.5.1 Grundlagen

Bei der Interpolation werden die einzelnen Datenpunkte durch Kurvenabschnitte in einer vorgegebenen Reihenfolge miteinander verbunden. In vielen Fällen sind die Datenpunkte in aufsteigender Reihenfolge nach den Abszissenwerten (x-Werte) geordnet. Dann lässt sich mit der Interpolationsfunktion für jede Stelle im Intervall der x-Werte der Daten ein zugehöriger y-Wert eindeutig berechnen. Interpolationsfunktionen lassen sich auch in mehrdimensionalen Datenräumen berechnen (z.B. als Kurve oder Fläche im dreidimensionalen Koordinatensystem).

Der Funktionstyp, der zur Interpolation verwendet werden soll, kann meist nicht aus den Daten abgeleitet werden. Kriterien für die Auswahl können sein:

- zu Grunde liegende Gesetzmäßigkeit
- geschlossene Funktionsgleichung für den gesamten Interpolationsbereich
- einfache mathematische Form
- Differenzierbarkeit
- Anpassung der Anschlussstellen (Wert, Steigung, Krümmung, ...)

Da alle Datenpunkte auf der Interpolationsfunktion liegen, eignet sie sich für die Fälle, bei denen die Datenpunkte mit großer Genauigkeit bestimmt werden konnten.

1.5.2 Lineare Interpolation

Bei der Linearen Interpolation werden die benachbarten Datenpunkte durch Lineare Funktionen (Geraden) verbunden. Dies ist eine mathematisch sehr einfache Form, die allerdings für jedes Intervall eine eigene Funktion erfordert. Bei jeder Berechnung muss daher zuerst das zugehörige Intervall bestimmt werden. An den Stützpunkten ergibt sich in der Regel ein Knick (nicht stetig differenzierbar).

Ein Interpolationswert y an der Stelle x zwischen den Stellen x_i und x_{i+1} mit den zugehörigen Werten y_i bzw. y_{i+1} lässt sich nach folgendem Ansatz berechnen:

$$\frac{(y - y_i)}{(y_{i+1} - y_i)} = \frac{(x - x_i)}{(x_{i+1} - x_i)}$$

Die Formel muss jetzt nach y umgestellt werden.

Zweipunkt-Kalibrierung

Bei der Zweipunkt-Kalibrierung wird zu zwei verschiedenen Merkmalswerten x_1 bzw. x_2 einer Probe der jeweils zugehörige Signalwert y_1 bzw. y_2 bestimmt. Durch zwei Punkte $P(x_1/y_1)$ und $Q(x_2/y_2)$ ist eine Gerade $y(x)=m \cdot x+b$ eindeutig bestimmt (Kalibrierfunktion).

In der Tabellenkalkulation können neben der Angabe der Formel im Diagramm (Trendlinie linear) auch die Tabellenfunktionen verwendet werden:

=STEIGUNG($y_1;y_2;x_1;x_2$)

=ACHSENABSCHNITT($y_1;y_2;x_1;x_2$)

Zur Berechnung des zu einem Signalwert y gehörenden Merkmalswert x ist die Umkehrung (Analysefunktion) $x(y)$ erforderlich.

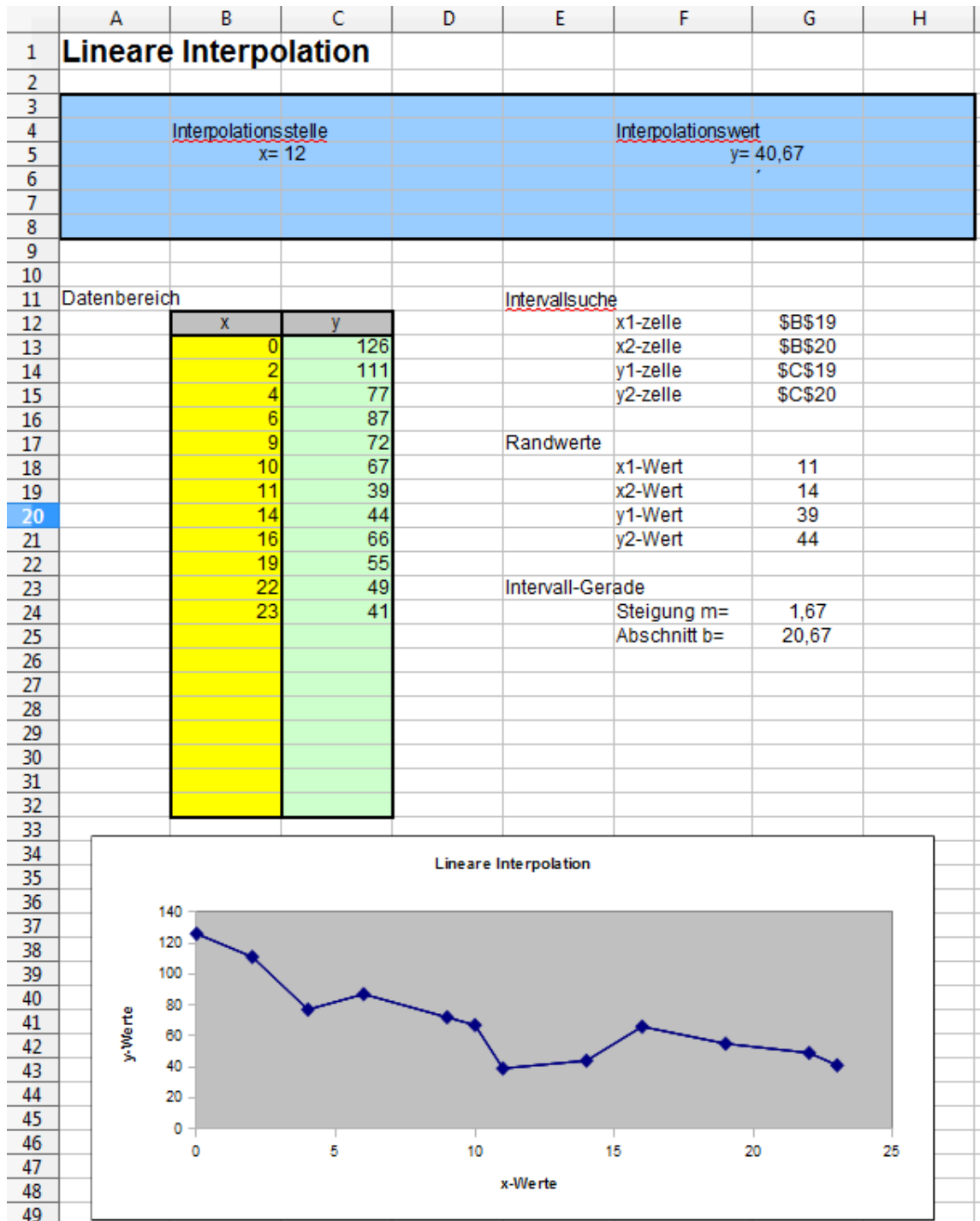
Lineare Mehrpunkt-Interpolation

In der Praxis wird eine Kalibrierfunktion häufig durch mehrere Punkte bestimmt. Dann ist für jedes Intervall zwischen zwei Punkten eine andere Geradengleichung erforderlich. Bei der Bestimmung des Merkmalswertes einer unbekannten Probe muss daher zuerst das zuständige Intervall ermittelt und dann über die zugehörige Geradengleichung der Wert berechnet werden. Außerhalb des Kalibrierbereichs sind keine sicheren Aussagen möglich.

Kurvenanpassung in EXCEL

Eine Lineare Interpolation wird angewendet, wenn die Datenpunkte nur geringe zufällige Schwankungen aufweisen.

In der Datei interlin.xlsx ist die Auswahl des zugehörigen Intervalls der Kalibrierfunktion und Berechnung des Interpolationswertes automatisiert. Die x-Werte (Merkmalswerte) müssen dazu eine (streng) aufsteigende Sortierung aufweisen. Wenn die y-Werte (Signalwerte) ebenfalls (streng) auf- oder absteigend sind, kann durch Vertauschung der beiden Reihen auch die Umkehrfunktion (Analysenfunktion) automatisiert werden.



Lösung mit Tabellenfunktionen

In der Tabellenkalkulation besteht die Schwierigkeit darin, das geeignete Intervall zwischen den Datenpunkten zu bestimmen.

Mit Tabellenfunktionen lässt sich die Aufgabe lösen. Die Datei interlin.xls kann sowohl von EXCEL als auch von CALC ausgeführt werden. Die x-Werte der Stützstellen haben den Bereichsnamen xwerte erhalten. Die Interpolationsstelle x wird in C5 eingegeben. Die Zelladressen mit der Startstelle x_i bzw. dem Startwert y_i wird bestimmt mit der Funktion:

=VERKETTEN("\$B\$";ZEILE(xwerte)+VERGLEICH(\$C\$5;xwerte)-1) bzw.
=VERKETTEN("\$B\$";ZEILE(xwerte)+VERGLEICH(\$C\$5;xwerte))

und entsprechend für die y-Werte

=VERKETTEN("\$C\$";ZEILE(xwerte)+VERGLEICH(\$C\$5;xwerte)-1) bzw.
=VERKETTEN("\$C\$";ZEILE(xwerte)+VERGLEICH(\$C\$5;xwerte))

Zu den zugehörigen Werten gelangt man mit

=INDIREKT(G12) und entsprechend G13, G14, G15

Jetzt lassen sich Steigung und Achsenabschnitt der Geraden berechnen und damit der gesuchte Interpolationswert y.

1.5.3 Polynom-Interpolation

Eine Polynomfunktion lässt sich durch folgende Funktionsgleichung darstellen:

$$y(x) = a_n \cdot x^n + \dots + a_3 \cdot x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0$$

Die Koeffizienten werden auch (beginnend mit a_0) mit den Buchstaben a, b, c, d, ... bezeichnet. Der höchste vorkommende Exponent ergibt den Grad (oder die Ordnung) des Polynoms.

Ein Polynom vom Grad n interpoliert n+1 Datenpunkte. Liegen weniger Punkte vor, sind die Werte der entsprechenden Koeffizienten gleich Null. Liegen mehr Datenpunkte vor, wird durch das Polynom im Allgemeinen nur eine Approximation an die Punkteverteilung erreicht.

1.5.4 Spline-Interpolation

Der Ursprung der Splinefunktionen liegt im Schiffsbau. Eine biegsame Latte (Spline) biegt sich zwischen gegebenen festen Punkten so, dass die Biegeenergie (Krümmungsintegral) ein Minimum erreicht. Damit lässt sich die Querschnittsform eines Schiffes modellieren. Mathematisch lässt sich diese Linie zwischen zwei Stützpunkten jeweils durch ein Polynom 3. Grades beschreiben. An den Anschlusspunkten stimmen Funktionswert, Steigung und Krümmung überein. Die Kurve wird auch durch die Bedingungen für Steigung und Krümmung am linken und am rechten Rand bestimmt. Beim natürlichen Spline geht die Randsteigung jeweils in eine Gerade über.

Spline-Interpolationen zeigen oft einen angenehm glatten Verlauf, erfordern aber für jedes Intervall eine eigene Funktionsgleichung und bei der Berechnung die Auswahl des jeweils zuständigen Intervalls.

Zur Berechnung der Polynomkoeffizienten der Spline-Abschnitte müssen relativ große Lineare Gleichungssystem (LGS) gelöst werden (mit pro Intervall jeweils vier Variablen).

Splines zeigen vor und nach einer Stufe in den gegebenen Datenpunkten ein unerwünschtes Überspringen. Dies wird beim Akima-Spline durch einen Verzicht auf übereinstimmende Krümmung an den Übergangsstellen vermieden.

Spline-Interpolationen werden in EXCEL nicht durch bereitgestellte Funktionen unterstützt. Die Option der Darstellung einer geglätteten Linie im Diagramm stellt eine andere, nicht dokumentierte Funktion dar.

1.6 Approximation

1.6.1 Grundlagen

Gleichwertige Bezeichnungen sind „Approximation“, „Ausgleich“, „Anpassung“, „Regression“ und „Fit“. Mit der Bezeichnung „Fit“ bezieht man sich oft auf ein iteratives Verfahren. In der Mathematik wird Approximation als Annäherung an eine bekannte Funktion verstanden, was in der Messtechnik jedoch nur selten vorkommt.

Der Begriff "Regression" bedeutet "Rückschritt" und ist das Überbleibsel einer Untersuchung der Körpergrößen von Vätern und Söhnen, die von F. Galton durchgeführt wurde. Große Väter haben zwar auch große Söhne, aber bei sehr großen Vätern sind die Söhne eher etwas kleiner als der Vater. Bei kleinen Vätern sind die Söhne eher etwas größer als der Vater. Es gibt also einen Rückschritt zum Mittelwert hin.

Bei der Approximation geht man davon aus, dass die vorliegenden Datenpunkte nicht exakt die Abhängigkeit zwischen den Variablen beschreiben. Es wird eine grafisch anschauliche oder mathematisch einfachere Funktion gesucht, die das zu Grunde liegende (oft nur vermutete) Verhalten zumindest näherungsweise beschreibt. Sind die zugehörigen Gesetzmäßigkeiten bekannt, wird man nur die entsprechenden Parameter der Funktion bestimmen, sind sie unbekannt, müssen Modellvorstellungen entwickelt oder andere mathematische Anpassungsbedingungen bestimmt werden.

Typische Anpassungsbedingungen sind:

- **Minimum der y-Abstandsquadrate**
Wegen des geringeren Rechenaufwandes betrachtet man nur die Abstände der Punkte von der Geraden in y-Richtung und fordert, dass die Summe der quadrierten Abstände möglichst klein (minimal) sein soll. Das Quadrieren erspart die Unterscheidung positiver und negativer Abstände und gewichtet große Abstände stärker.
Prinzip: Die Summe der quadrierten y-Abstände soll minimal sein.
- **Minimum der Summe der Abstandsbeträge**
(kann nur schrittweise angenähert werden, iterativ)
- **kleinster Maximalabstand**
(kann nur schrittweise angenähert werden, iterativ)

Es sieht so aus, als ob die Kurve von den Datenpunkten angezogen wird. Durch eine unterschiedliche Gewichtung kann man den einzelnen Datenpunkten unterschiedliche „Kräfte“ zuordnen. Die wird dann sinnvoll sein, wenn einzelne Datenpunkte mit unterschiedlicher Messunsicherheit bestimmt worden sind.

Bei unbedeutender Messunsicherheit ist eine Interpolation sinnvoll. Bei geringer Anzahl von Stützpunkten geht die Approximation in die Interpolation über (wenn die Anzahl der Stützstellen gleich der Anzahl der Funktionsparameter ist).

1.6.2 Direkte Verfahren

Bei den direkten Verfahren gibt es eine Formel oder Rechenvorschrift mit der aus den gegebenen Daten der Stützpunkte auf algebraischem Weg die Parameter der gesuchten Funktion bestimmt werden können.

- **Gerade**
Für die „Ausgleichsgerade“ sind die Formeln unter 1.1.1 angegeben.
- **Polynom**
Die Polynomkoeffizienten lassen sich über die Lösung eines Linearen Gleichungssystems (LGS) bestimmen. Die Formeln sind unter 3.2 angegeben. Ein EXCEL-Beispiel findet man

unter „polyprox.xlsx“ bzw. „polyproxa.xlsx“.

- **Spline**
Spline-Approximationen werden in EXCEL nicht durch bereitgestellte Funktionen unterstützt.
Beispiele findet man (11/2013) unter <http://www.excelformeln.de/tips.html> .

1.6.3 Linearisierung

Eine Anpassung über die Linearisierung kann überall dort vorgenommen werden, wo in einem streng monotonen Funktionsabschnitt die Umkehrfunktion berechnet werden kann.

- **Potenzfunktion**
Durch Logarithmieren werden die x- und die y-Werte in eine lineare Abhängigkeit transformiert, dazu eine Lineare Funktion (durch den Ursprung) bestimmt und anschließend mit der Exponentialfunktion in die ursprüngliche Skalierung zurück transformiert.
- **Exponentialfunktion**
Durch Logarithmieren werden die y-Werte in eine lineare Abhängigkeit transformiert, dazu eine Lineare Funktion bestimmt und anschließend mit der Exponentialfunktion in die ursprüngliche Skalierung zurück transformiert.
- **Logarithmusfunktion**
Durch die zugehörige Exponentialfunktion werden die y-Werte in eine lineare Abhängigkeit transformiert, dazu eine Lineare Funktion bestimmt und anschließend mit der Exponentialfunktion in die ursprüngliche Skalierung zurück transformiert.

1.6.4 Iterative Verfahren

Bei einer Iteration werden ausgehend von geeigneten Startwerten die Parameter der Anpassungsfunktion bei jedem Schritt so verändert, dass sich die Qualität der Anpassung verbessert. Bei ausreichender Nähe zur gesuchten Lösung wird das Verfahren abgebrochen. Bei ungünstigen Startparametern kann das Verfahren auch divergieren.

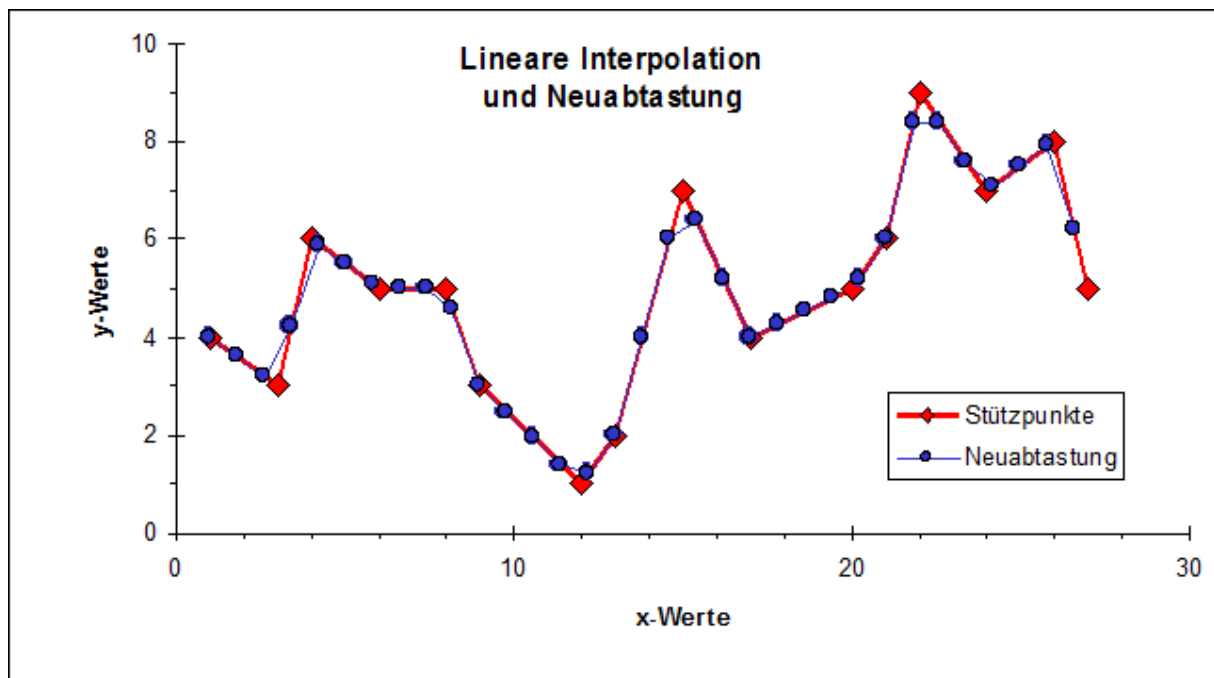
- **Freier Funktionsansatz**
Es können beliebige Funktions-Ansätze gewählt werden. Die Berechnung kann oft nur mit speziellen Rechenprogrammen (Algorithmen) durchgeführt werden.

Ein häufig anzutreffender Algorithmus ist das Verfahren nach „Marquardt-Levenberg“.

2 Neuabtastung

Sollen Datensätze miteinander verglichen oder verrechnet werden, sollten sie identische Abszissenwerte (x-Werte) besitzen. Häufig ergibt sich auch die Forderung nach einer äquidistanten Abtastung (gleicher Abstand der x-Werte). Für den originalen Datensatz wird eine Interpolationsfunktion (eventuell auch eine Anpassungsfunktion) bestimmt und dies dann im erforderlichen Raster abgetastet. In vielen Fällen reicht es aus, hier eine lineare Interpolation zu verwenden. Die Anzahl der neuen Datenpunkte kann jetzt beliebig groß gewählt werden.

Bei der Neuabtastung werden die Maximas und Minimas in der Regel nicht genau getroffen, so dass die neue Reihe einen kleineren Variationsbereich besitzt. Um nicht wichtige Informationen zu verlieren, sollte bei der Neuabtastung ein Vergleich der Extremas der ursprünglichen und der abgetasteten Reihe vorgenommen werden.



3 Kurvenanpassung mit Standardfunktionen in EXCEL

3.1 Lineare Funktion

(andere Bezeichnungen: lineare Trendlinie, Lineare Regression, Ausgleichsgerade)

Funktionsgleichung $y(x) = m \cdot x + b$ (oder auch $y(x) = b \cdot x + a$)

Wichtige Eigenschaften sind:

- Steigung m und Achsenabschnitt b werden so bestimmt, dass die Summe der quadrierten y -Abstände ein Minimum erreicht. Hierfür gibt es direkte (geschlossene) Formeln. Die Parameter m und b können auch mit Tabellenfunktionen berechnet werden.

=STEIGUNG(ywerte;xwerte)

=ACHSENABSCHNITT(ywerte;xwerte)

- Zur Beschreibung der Güte der Anpassung können Bestimmtheitsmaß R^2 und Korrelationskoeffizient R berechnet werden. Bei $R=\pm 1$ bzw. $R^2=1$ liegen alle Punkte auf der Geraden. R kann Werte zwischen -1 und $+1$ annehmen. Das Vorzeichen entspricht dem Vorzeichen der Steigung. Bei den Befehlen darf hier die Reihenfolge der Bereiche auch vertauscht werden.

=BESTIMMTHEITSMAB(ywerte;xwerte)

=KORREL(ywerte;xwerte)

Pearsonscher Korrelationskoeffizient

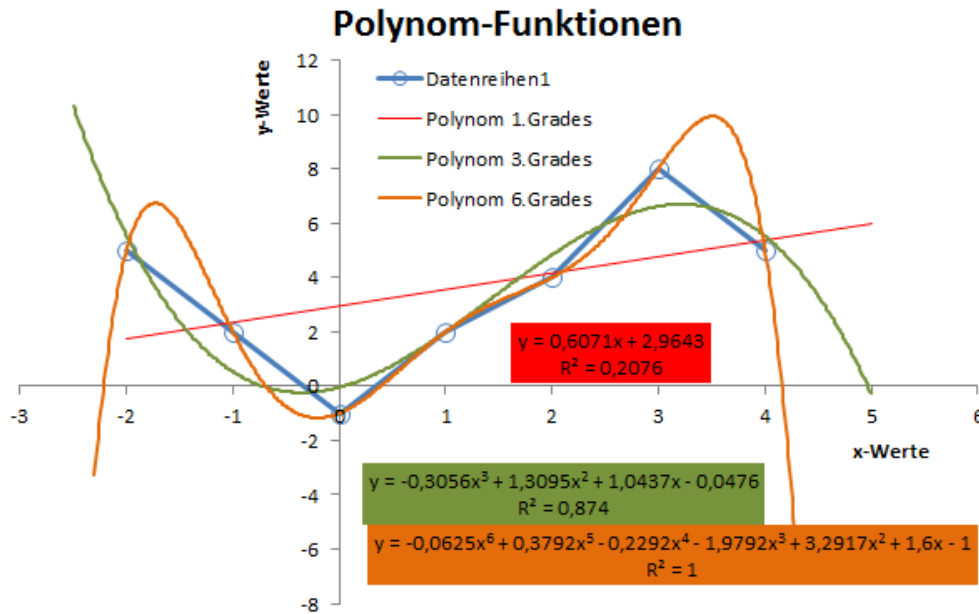
- Der Sonderfall, dass alle Punkte auf einer waagrechten Geraden liegen wird in EXCEL nicht berechnet. Es müsste sich dabei $R=R^2=0$ ergeben, da die y -Werte nicht von den x -Werten abhängen.
- Im amerikanischen Sprachgebrauch heißt das Bestimmtheitsmaß „r-squared“.
- Für andere Fragestellungen gibt es weitere Korrelationsmaße (z.B. Spearman, Kendall).
- Im AddIn „Analysefunktionen“ können mit der Funktion „Regression“ weitere Kenngrößen berechnet werden.

3.2 Polynomfunktion

Funktionsgleichung $y(x) = a_n \cdot x^n + \dots + a_3 \cdot x^3 + a_2 \cdot x^2 + a_1 \cdot x + a_0$

Wichtige Eigenschaften sind:

- Ein Polynom n -ten Grades interpoliert $n+1$ Punkte
- Ein Polynom n -ten Grades hat bis zu n Nullstellen
- Ein Polynom n -ten Grades hat bis zu $n-1$ Extremas
- Ein Polynom n -ten Grades hat bis zu $n-2$ Wendepunkte
- Mit Polynomen lassen sich viele Rechenoperationen einfach ausführen. Sie besitzen keine Sprungstellen und sind überall beliebig oft differenzierbar und integrierbar.
- Polynome eignen sich nur in Sonderfällen zur Interpolation, da sie stark zu Überschwängern neigen.
- Das Ausgleichspolynom kann direkt über die Lösung eines Linearen Gleichungssystems berechnet werden.



Rechenverfahren

Aus den Messdaten (x-Werte, y-Werte) müssen die untenstehende Matrix A und der Konstantenvektor c berechnet werden. Dabei ist n der Grad des Polynoms. Die Matrix A wird dann numerisch invertiert und mit dem Vektor c multipliziert. Der Ergebnisvektor b enthält dann die Koeffizienten des Ausgleichspolynoms.

Aus $A \cdot b = c$ wird nach der Umstellung $b = A^{-1} \cdot c$

Matrix A:

n	$\Sigma(x_i)$	$\Sigma(x_i^2)$	$\Sigma(x_i^3)$...	$\Sigma(x_i^n)$	$\Sigma(y_i)$
$\Sigma(x_i)$	$\Sigma(x_i^2)$	$\Sigma(x_i^3)$	$\Sigma(x_i^4)$...	$\Sigma(x_i^{n+1})$	$\Sigma(y_i \cdot x_i)$
$\Sigma(x_i^2)$	$\Sigma(x_i^3)$	$\Sigma(x_i^4)$	$\Sigma(x_i^5)$...	$\Sigma(x_i^{n+2})$	$\Sigma(y_i \cdot x_i^2)$
$\Sigma(x_i^3)$	$\Sigma(x_i^4)$	$\Sigma(x_i^5)$	$\Sigma(x_i^6)$...	$\Sigma(x_i^{n+3})$	$\Sigma(y_i \cdot x_i^3)$
...
$\Sigma(x_i^n)$	$\Sigma(x_i^{n+1})$	$\Sigma(x_i^{n+2})$	$\Sigma(x_i^{n+3})$...	$\Sigma(x_i^{2n})$	$\Sigma(y_i \cdot x_i^n)$

Konstantenvektor c:

3.3 Potenzfunktion

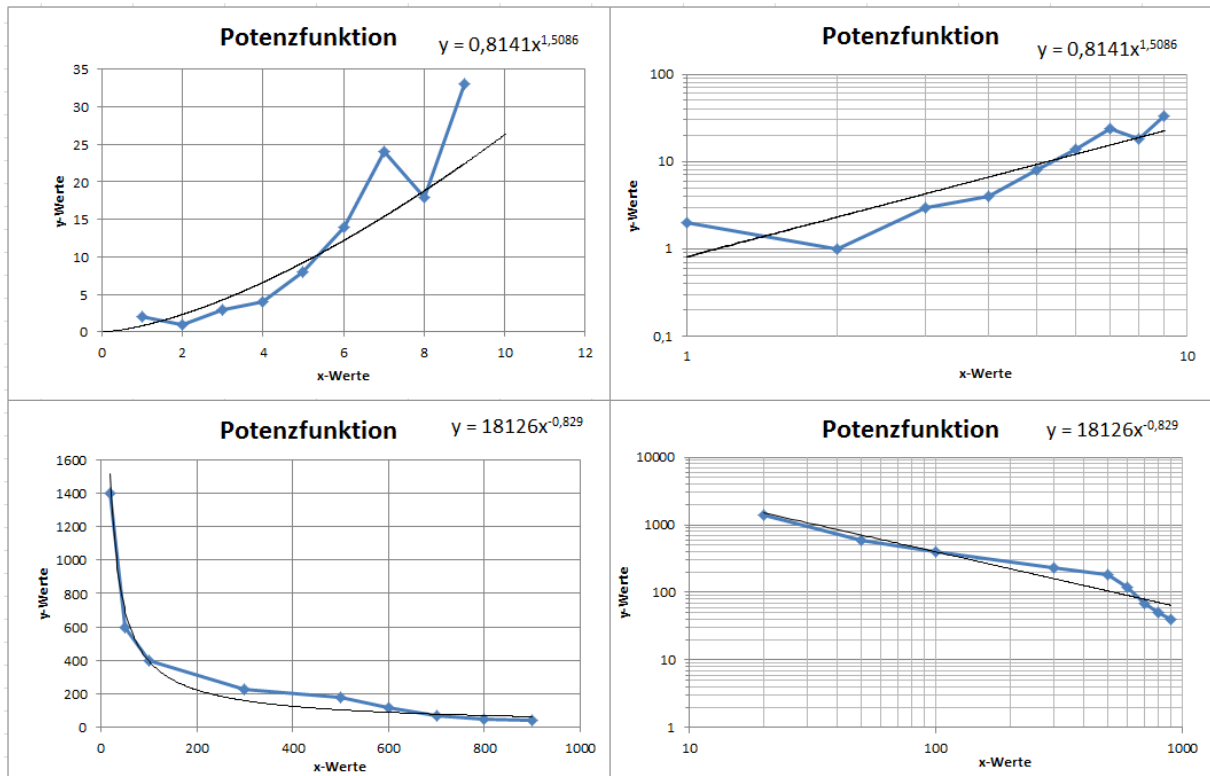
Funktionsgleichung $y(x) = a \cdot x^b$

Wichtige Eigenschaften sind:

- Potenzfunktionen mit $b \geq 0$ gehen immer durch den Ursprung.
- Potenzfunktionen mit $b < 0$ haben eine Unstetigkeit bei $x=0$.
- Alle x-Werte und alle y-Werte müssen in EXCEL positiv sein. Mathematisch sind teilweise auch andere Kombinationen möglich.
- Die Polynomfunktion wird in EXCEL so berechnet:
 - Logarithmieren der x- und y-Werte
 - Bestimmung einer Ausgleichsgeraden
 - Rücktransformation mit Exponentialfunktion

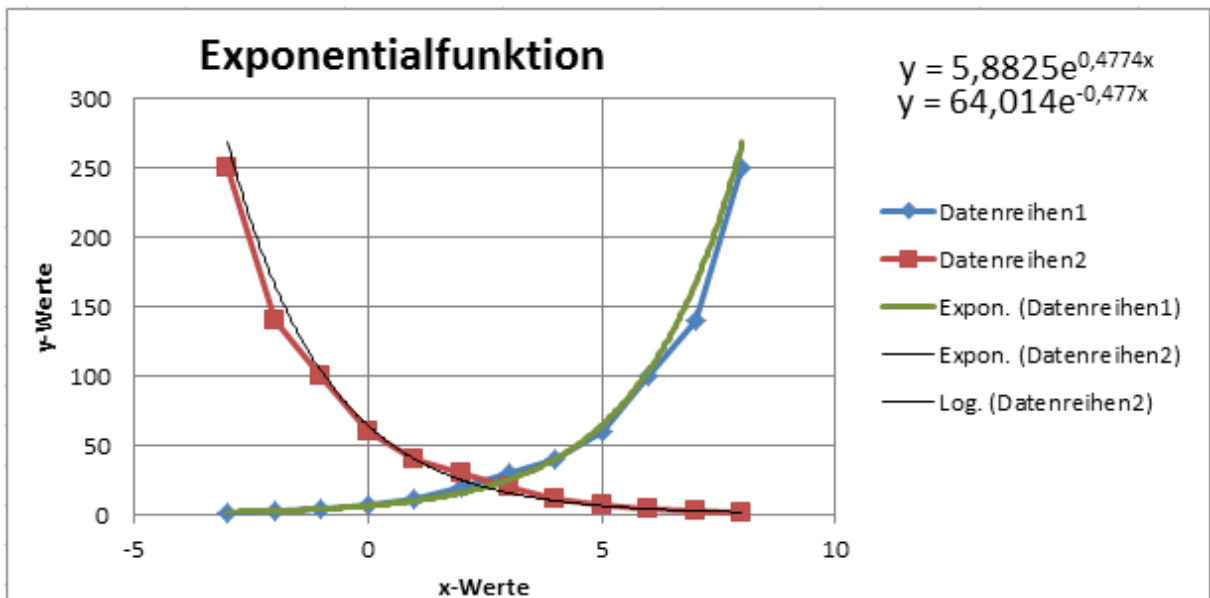
Kurvenanpassung in EXCEL

- Durch die Linearisierung werden kleine Werte viel stärker gewichtet als große. R und R² sind nur für Vergleiche brauchbar.
- Bei doppelt-logarithmischer Skalierung der Diagrammachsen ergibt sich eine Gerade.



3.4 Exponentialfunktion

Funktionsgleichung $y(x) = a \cdot e^{b \cdot x}$

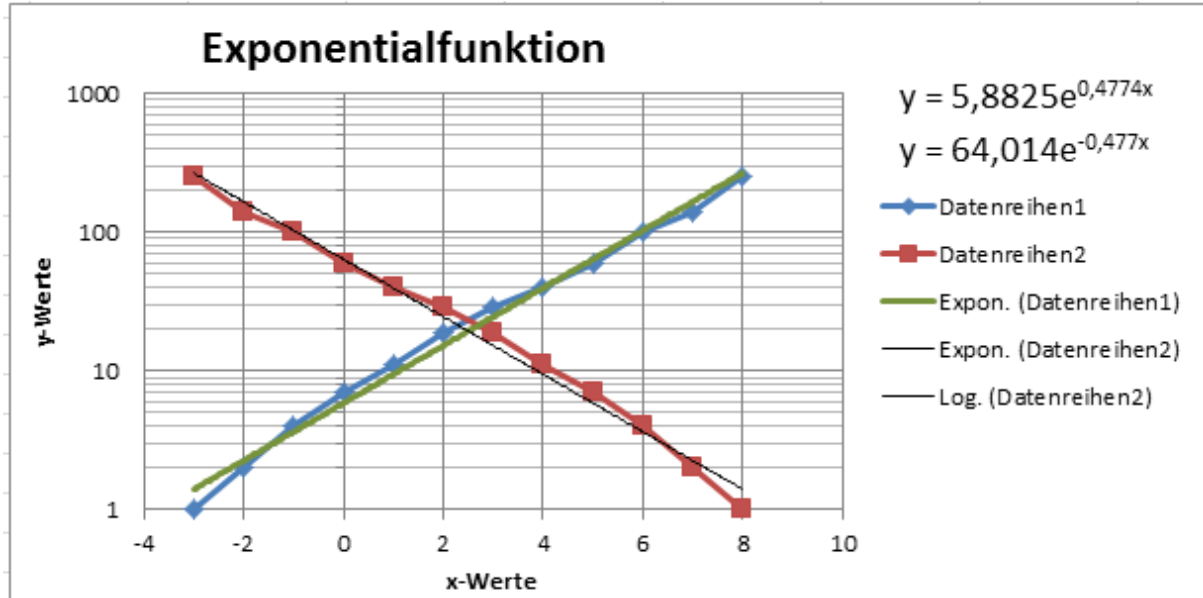


Wichtige Eigenschaften sind:

- Es wird nur die e-Funktion ($e=2,71828\dots$) angeboten. Eine Umrechnung in andere Basen ist möglich.
Bei der Darstellung zur Basis c gilt $y = a \cdot c^x = a \cdot e^{\ln(c) \cdot x}$ (mit $b=\ln(c)$)
- Die y-Werte müssen positiv sein. Mathematisch wären auch rein negative y-Werte möglich.
- Die Exponentialfunktion wird in EXCEL so berechnet:

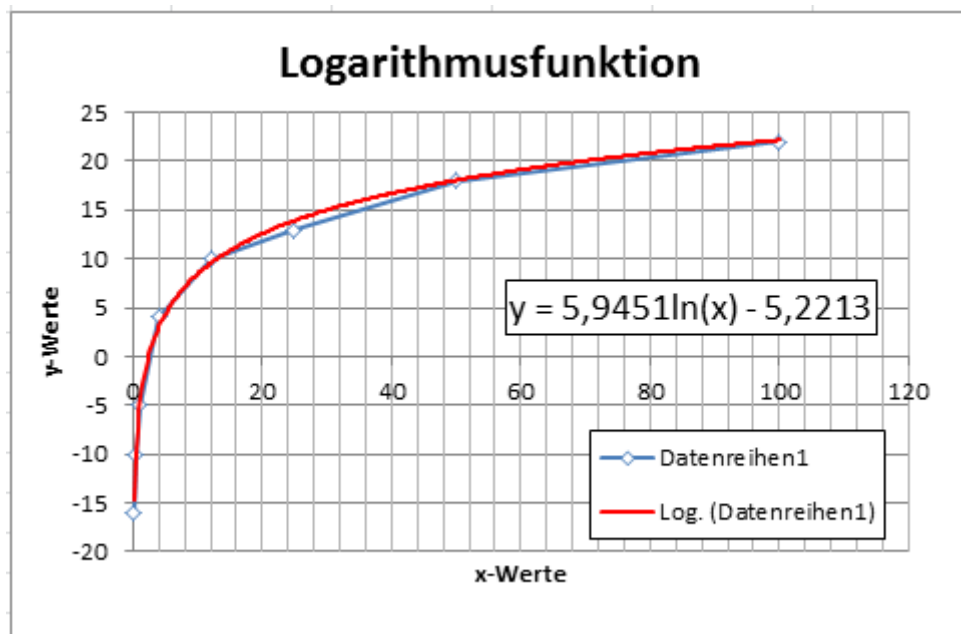
Kurvenanpassung in EXCEL

- Logarithmieren der y-Werte
- Bestimmung einer Ausgleichsgeraden
- Rücktransformation mit Exponentialfunktion
- Durch die Linearisierung werden kleine Werte stärker gewichtet als große. R und R^2 sind nur für Vergleiche brauchbar.
- Bei logarithmischer Skalierung der y-Diagrammachse ergibt sich eine Gerade.



3.5 Logarithmusfunktion

Funktionsgleichung $y(x) = a \cdot \ln(x) + c$ (ln: natürlicher Logarithmus, Basis $b=e$)

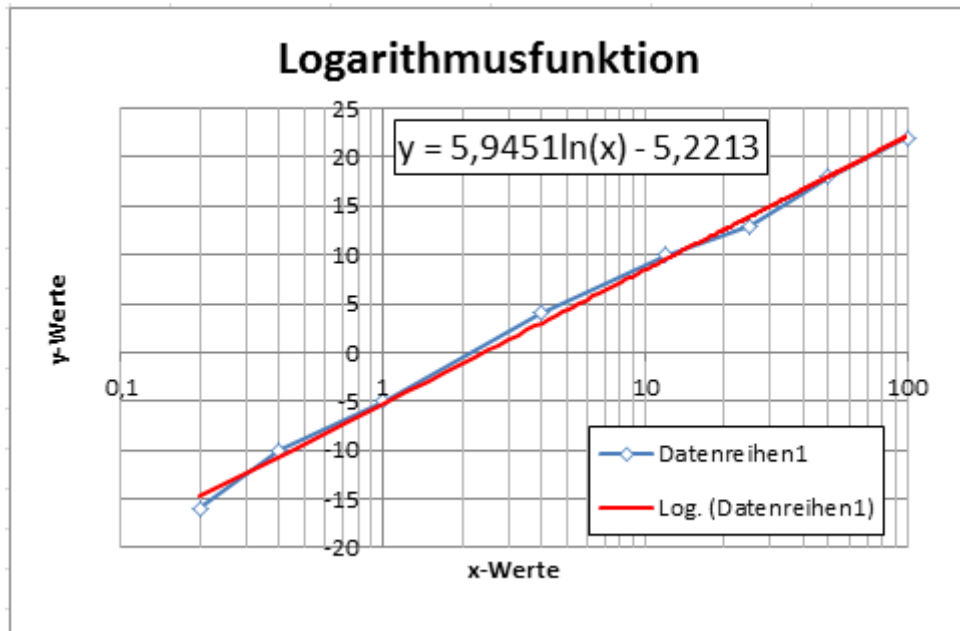


Wichtige Eigenschaften sind:

- Es wird nur der natürliche Logarithmus \ln angeboten. Eine Umrechnung in andere Basen ist möglich. Für die Umrechnung in den Logarithmus zur Basis b gilt:
 $a \cdot \ln(x) + c = a \cdot \ln(b) \cdot \log_b(x) + c$
- Die x -Werte müssen positiv sein.
- Die Exponentialfunktion wird in EXCEL so berechnet:
 - Logarithmieren der x -Werte

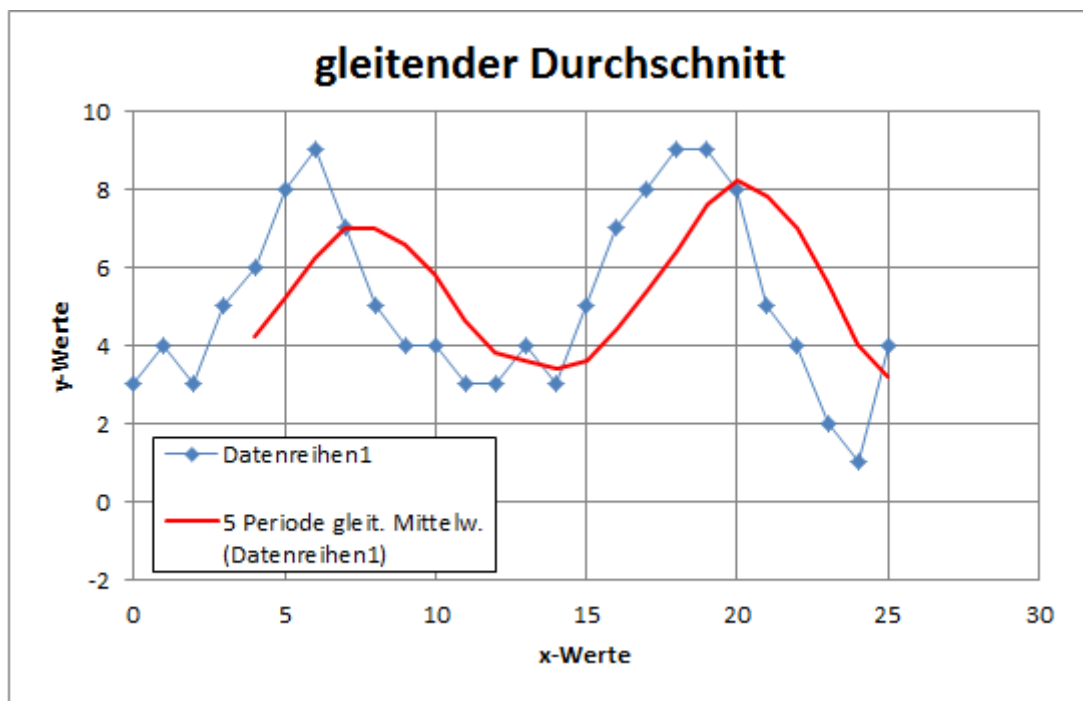
Kurvenanpassung in EXCEL

- Bestimmung einer Ausgleichsgeraden
- Rücktransformation mit Exponentialfunktion
- Durch die Linearisierung werden große Werte stärker gewichtet als kleine. R und R² sind nur für Vergleiche brauchbar.
- Bei logarithmischer Skalierung der x-Diagrammachse ergibt sich eine Gerade.



3.6 Gleitender Durchschnitt

Diese Funktion eignet sich in der Regel nicht für die Weiterverarbeitung bereits erfasster Messdaten. Für die Glättung werden nur zeitlich zurückliegende Daten verwendet (kausales Verfahren). Dies führt dazu, dass die geglättete Kurve gegenüber den originalen Daten verschoben ist. Zur Nachbearbeitung von Messdaten (off line) sollte daher besser ein symmetrisches Glättungsfilter (akausal) verwendet werden.



4 Kurvenanpassung mit dem SOLVER

4.1 Iterative Verfahren (SOLVER)

Bei einem iterativen Verfahren wird eine Lösung durch schrittweise Annäherung gefunden. Der Vorteil liegt in der Verwendung einfacher Rechenoperationen und darin, dass der Funktionszusammenhang oft nicht als Funktionsgleichung bekannt sein muss (z.B. bei Integral, Differentialgleichung). Ein Nachteil kann sein, dass man immer nur Näherungslösungen erhält und eine Fehlerabschätzung oft schwieriger ist als die eigentliche Berechnung.

Der Solver in EXCEL ist ein solches Rechenverfahren. Die Ausgangswerte für eine mögliche Lösung werden systematisch variiert und der Abstand des Rechenergebnisses zu den durch das Problem gegebenen Bedingungen bewertet (z.B. dass eine in der Rechnung vorkommende Größe möglichst klein wird). Das Verfahren bricht ab, wenn dieser Abstand zu den Bedingungen einen vorgebbaren Wert unterschreitet oder wenn eine vorgebbare Anzahl von Iterationsschritten erreicht ist.

4.2 Durchführung der Berechnung

Folgende Schritte sind in der Regel durchzuführen:

- Festlegung des Funktionstyps
- Aufstellen der formalen Funktionsgleichung
- Festlegung der variierbaren Parameter
- Festlegung der Minimierungsbedingung
- Optimierung der Parameter

Das verwendete Rechenverfahren ist in EXCEL nicht offengelegt. Vermutlich wird das Iterationsverfahren nach Marquardt-Levenberg angewandt.

Als Minimierungsbedingung verlangt man meist, dass die Summe der quadrierten Abstände der Datenpunkte von der Ausgleichsfunktion minimal sein soll.

Dazu wird an jeder Stelle x_i der quadrierte Abstand $(y_i - f(x_i))^2$ berechnet und von allen Ergebnissen die Summe gebildet. Diese Summe soll möglichst klein werden.

Bei Verwendung der relativen Abstände wird die aus der Statistik bekannte Größe Chi-Quadrat verwendet. Dazu wird der quadrierte Abstand durch den erwarteten Funktionswert dividiert. Auch hier soll die Summe dieser Werte möglichst klein werden.

$$CHI2 = \sum [(y_i - f(x_i))^2 / f(x_i)]$$

Welches Abstandsmaß geeigneter ist, muss von der Problemstellung her entschieden werden. Die relativen Abstände gewichten die Abweichungen bei kleinen Funktionswerten stärker.

4.3 Beispiel aus der Spektroskopie

4.3.1 Ansatz

Spektrallinien lassen sich meist durch eine Gauß-Funktion beschreiben. Rauschen im Messprozess lässt die einzelnen Werte streuen und führt zu einem Untergrund, der im einfachsten Fall durch eine Konstante beschrieben werden kann. Der geeignete Funktionsansatz ist damit die Summe aus Gauß-Funktion und Konstante. Die Funktionsgleichung lässt sich so beschreiben:

$$y(x) = a \cdot \exp(-((x-b)^2 / (2 \cdot s^2)) + d$$

in EXCEL muss eine zusätzliche Klammer verwendet werden:

$$y(x) = a \cdot \exp(-(((x-b)^2) / (2 \cdot s^2))) + m \cdot x + d$$

mit folgender Zuordnung in der Formel:

- a: Maximalwert
- b: Lage des Maximums
- s: Breite der Verteilung (Standardabweichung)
- m: Steigung
- d: Konstante (Untergrund)

Alternativ kann die Formel für die Normalverteilung verwendet werden, bei der der Parameter a dann bereits der Peakfläche ohne Basisanteil entspricht.

$$= \text{NORMVERT}(x; b; s; \text{falsch}) + m \cdot x + d$$

Die Parameter a bis d sind nun so zu variieren, dass die Summe der absoluten Abstandsquadrate minimal wird. Der logische Parameter „falsch“ schaltet zwischen Dichte- und Verteilungsfunktion um.

4.3.2 Realisierung der Anpassung

Die x-Werte und y-Werte sind als Messdaten vorgegeben. Die Parameter a bis d werden zuerst geschätzt. Bei schlechter Schätzung kann das Iterationsverfahren divergieren.

Geschätzte Startwerte

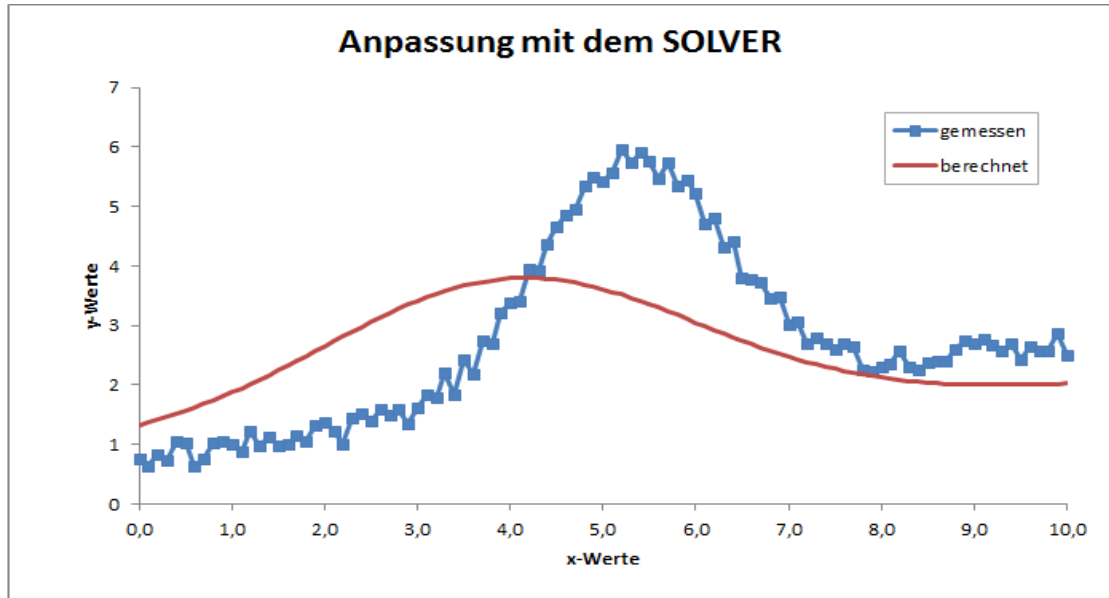
a=	12
b=	4
s=	2
m=	0,1
d=	1

Der SOLVER ist nicht standardmäßig installiert, sondern wird hinzugefügt durch:

Datei - Optionen - Add-Ins - Solver - verwalten - gehe zu - Haken setzen - OK

Er wird dann über die Registerkarte „Daten“ aufgerufen. Zielzelle und veränderbare Zellen lassen sich durch Markierung eintragen. Das Verfahren kann jetzt gestartet und die Lösung übernommen werden. Es gibt keine weiteren Nebenbedingungen. Die voreingestellten Optionen waren hier geeignet. Nach jeder Änderung in der Tabelle muss der SOLVER neu gestartet werden.

Kurvenanpassung in EXCEL



Solver-Parameter

Ziel festlegen:

Bis: ☐ Max. ☒ Min. ☐ Wert:

Durch Ändern von Variablenzellen:

Unterliegt den Nebenbedingungen:

Hinzufügen
Ändern
Löschen
Alles zurücksetzen
Laden/Speichern

☒ Nicht eingeschränkte Variablen als nicht-negativ festlegen

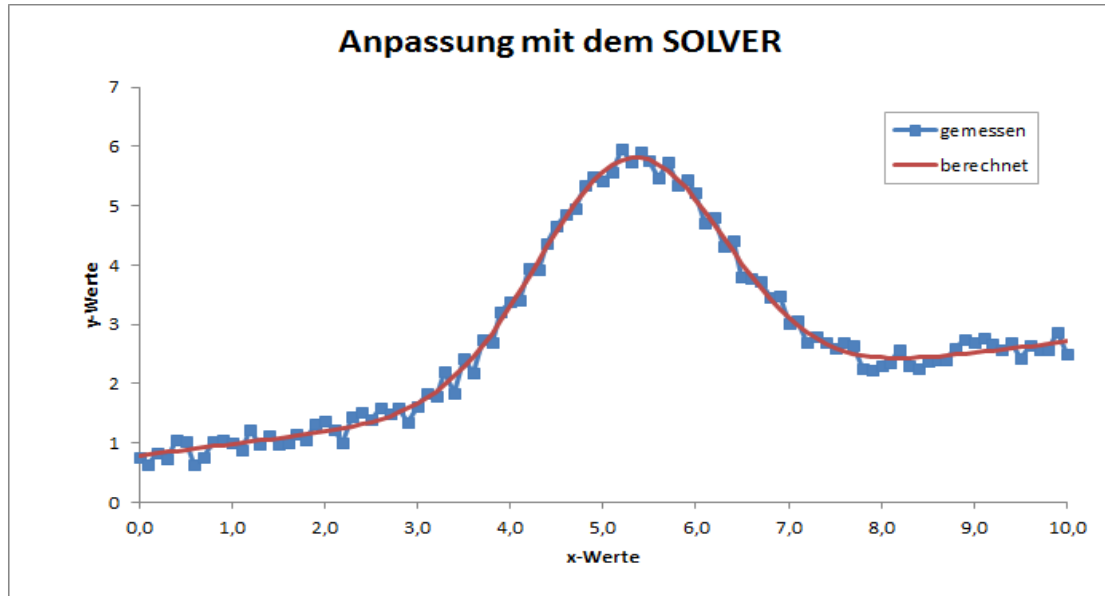
Lösungsmethode auswählen:

Lösungsmethode
Wählen Sie das GRG-Nichtlinear-Modul für Solver-Probleme, die kontinuierlich nichtlinear sind. Wählen Sie das LP Simplex-Modul für lineare Solver-Probleme und das EA-Modul für Solver-Probleme, die nicht kontinuierlich sind.

Berechnete Parameterwerte

a= 10,1409916
b= 5,30456662
s= 1,01060283
m= 0,19201845
d= 0,79126601

Kurvenanpassung in EXCEL



Kurvenanpassung mit dem SOLVER				
Parameter	a=	10,1409916		
	b=	5,30456662		
	s=	1,01060283		
	m=	0,19201845		
	d=	0,79126601		
		Summe:	2,015	
einzelner Peak mit Basislinie				
	x	y(gemessen)	y(berechnet)	quadrat
	0,0	0,748	0,791	0,002
	0,1	0,639	0,810	0,029
	0,2	0,838	0,830	0,000
	0,3	0,743	0,849	0,011
	0,4	1,053	0,868	0,034
	0,5	1,022	0,887	0,018

Erweiterung der Aufgabe:

In der Spektroskopie oder Chromatografie sind häufig zwei oder mehr Spektrallinien überlagert, so dass man ihre Lage, Breite oder Fläche nicht gesondert bestimmen kann. Simulieren Sie die entsprechenden Daten und passen Sie sie mit einer Funktion aus zwei Gauß-Funktionen plus linearem Untergrund an.

4.4 Weiteres SOLVER-Beispiel: Potentiometrische Titration

Bei einer potentiometrischen Titration wird zur vorgelegten Substanz (Titrand) schrittweise, langsam eine andere Substanz (Titrator) zugegeben. Bei der einsetzenden Reaktion ändert sich der pH-Wert, der nach jeder Zugabe mit einer pH-Elektrode gemessen wird. Das Messsignal bzw. der pH-Wert ergibt in Abhängigkeit vom zudosierten Volumen eine Stufen-Funktion, die durch eine sog. „Logistische Funktion“ (Schwanenhalskurve) beschrieben werden kann. Der gesuchte Äquivalenzpunkt der Reaktion lässt sich aus der Lage des Wendepunktes der Kurve bestimmen. Dieser befindet sich im Allgemeinen zwischen zwei Messpunkten. Oft sind auch die Volumen-Intervalle nicht gleich groß.

Eine verallgemeinerte Form der „Logistischen Funktion“ (unsymmetrisch zum Wendepunkt) wird durch folgende Funktionsgleichung beschrieben:

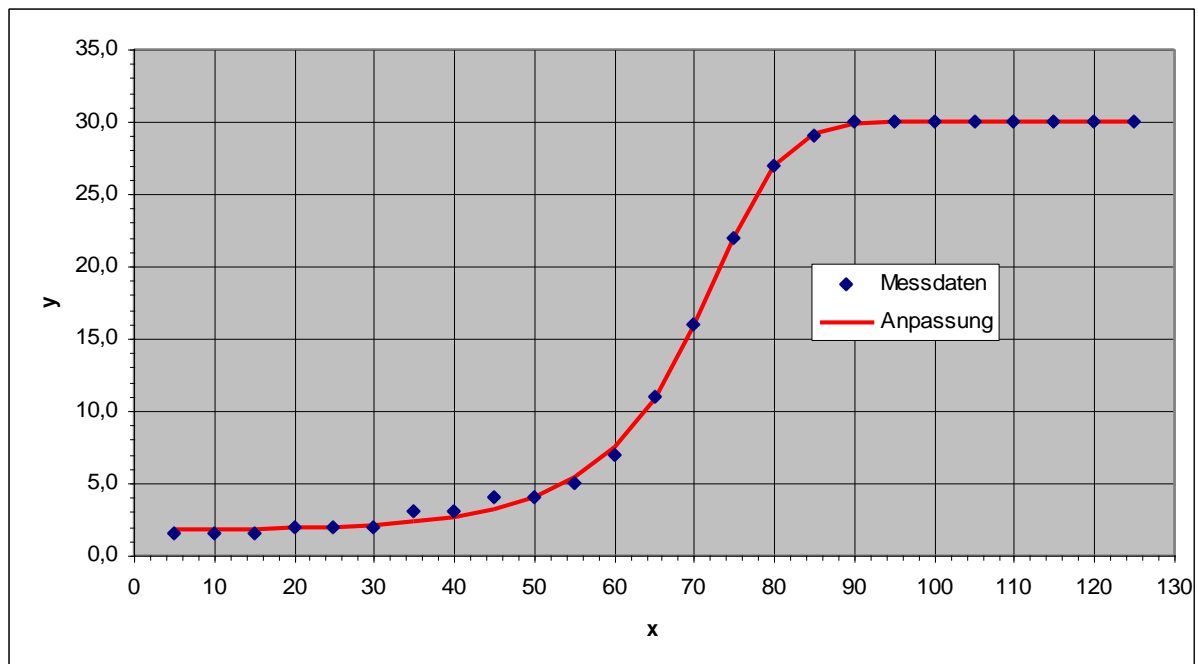
$$Y = A + \frac{C}{(1 + Te^{-B(x-M)})^{\frac{1}{T}}}$$

A=	1,774 (Basis)
C=	28,241 (Sprunghöhe)
T=	3,164 (Symmetrie)
B=	0,293 (Steilheit)
M=	73,205 (Wendestelle)

Durch die verschiedenen Parameter lassen sich die Eigenschaften der Kurve an die erfassten Messdaten anpassen. Die Symmetrie lässt dabei verschiedene Krümmungen an den Übergängen zu. Die für die Auswertung interessierende Wendestelle wird durch den Parameter M bestimmt.

Der SOLVER minimiert die Summe der quadrierten Abweichungen.

Schreibweise der Berechnungsformel in EXCEL: =A+C/((1+T*EXP(-B*(C11-M)))^(1/T))



Der SOLVER wird in EXCEL installiert mit:

Datei - Optionen - Add-Ins - Solver - verwalten -gehe zu - Haken setzen - OK

5 Übungsblätter

5.1 Lineare Kalibrierfunktion

Beim Kalibrieren wird der Zusammenhang zwischen einer messbaren Signalgröße (z.B. Extinktion) und der gesuchten Merkmalsgröße einer Probe (z.B. Massenkonzentration) festgestellt und mathematisch beschrieben.

Im vorliegenden Beispiel sind 10 Datenpunkte einer Kalibrierung gegeben. Für 10 Lösungen bekannter Massenkonzentration (Standards) wurden mit einem Photometer die zugehörigen Werte der Extinktion gemessen.

Nr.	β in mg/L	E in 1
1	40	0,17
2	60	0,20
3	80	0,31
4	100	0,37
5	120	0,45
6	140	0,49
7	160	0,56
8	180	0,61
9	200	0,66
10	220	0,69

Aufgabe a) Bestimmen Sie die Gleichung der Linearen Ausgleichsfunktion (Ausgleichsgerade). Geben Sie die Parameter mit vier signifikanten Stellen an.

Mit der Gleichung der Ausgleichsgeraden lassen sich Werte schätzen, die nicht gemessen wurden. Dies ist in zwei Richtungen möglich:

Die **Kalibrierfunktion** beschreibt in unserem Beispiel die Abhängigkeit der Extinktion von der Massenkonzentration. Die Schreibweise $E=E(\beta)$ besagt, dass E eine Funktion von β ist.

Aufgabe b) Welche Extinktion ist für eine Massenkonzentration von $b=125\text{mg/L}$ zu erwarten? Erstellen Sie im Tabellenblatt einen Bereich, in dem nach Eingabe der Massenkonzentration automatisch die zugehörige Extinktion berechnet wird (auch wenn die Kalibrierdaten geändert werden).

Die **Analysenfunktion** dreht die Problemstellung um beschreibt in unserem Beispiel die Abhängigkeit der Massenkonzentration von der Extinktion. Die Schreibweise $\beta=\beta(E)$ besagt, dass β eine Funktion von E ist.

Aufgabe c) Geben Sie Gleichung der Analysenfunktion als Geradengleichung an. Welche Massenkonzentration ist für eine Extinktion von $E=0,63$ zu erwarten? Erstellen Sie im Tabellenblatt einen Bereich, in dem nach Eingabe der Extinktion automatisch die zugehörige Massenkonzentration berechnet wird (auch wenn die Kalibrierdaten geändert werden).

In den meisten Fällen weichen die gemessenen Datenpunkte von der Ausgleichskurve ab. Ein Maß für die Stärke der Abweichung ist das Bestimmtheitsmaß R^2 oder die Korrelation R. Mehr Information steckt in der Darstellung der einzelnen Abweichungen (in y-Richtung). Diese Restabweichungen nennt man auch Residuen (Singular: Residuum).

Aufgabe d) Berechnen Sie die Residuen und stellen Sie diese in einem Säulendiagramm dar. Berechnen Sie auch probeweise die Summe der Residuen. Eine wichtige Größe in der Statistik ist die Summe der quadrierten Residuen.

Zusatz: Vergleich mit quadratischem Ausgleich

Die grafische Darstellung der Residuen zeigt manchmal eine Systematik. Dann sind die Abweichungen wohl nicht ganz zufällig. Statt einer Linearen Anpassung führt dann eine andere mathematische Funktion zu kleineren Abweichungen.

Aufgabe e) Kopieren Sie Ihr Diagramm mit der Linearen Ausgleichsfunktion und setzen Sie dort eine quadratische Ausgleichsfunktion (Parabel) ein. Vergleichen Sie die beiden Werte für das Bestimmtheitsmaß.

Aufgabe f) Wie lässt sich bei den Residuen erkennen, dass eine Parabel besser geeignet ist?

Aufgabe g) Führen Sie wie in b) eine automatische Berechnung mit der Kalibrierfunktion durch.

Aufgabe h) Welche Probleme treten bei der Analysenfunktion auf? Wie lassen sich diese lösen?

6 Hinweise

6.1 Dokumentenstruktur

Das Dokument liegt als zusammenhängende WORD-Datei „kurvenanpassung.docx“ vor. Für den Einsatz im Unterricht und im CULM sind die einzelnen Themen als pdf getrennt. Daher wiederholen sich auch einige Informationen.

Als Schüler-Übung bietet sich an, die Texte-Abschnitte ohne Bilder als docx-Datei auszugeben und sie durch Bildschirm-Kopien der zugehörigen EXCEL-Übungen ergänzen zu lassen.

6.2 Text-Dateien

Kurvenanpassung.docx	kompletter Text mit Bildern
Glättung.pdf	
Ausgleichsgerade.pdf	
Interpolation.pdf	
Approximation.pdf	
Standardfunktionen.pdf	
Solver.pdf	

6.3 EXCEL-Dateien

Glaettbeispiele.xlsx	Daten für verschiedenen Glättungsverfahren
Glaettbeispiele0.xlsx	Lösung dazu
Average.xlsx	Averaging
Interlin.xlsx	Lineare Interpolation an einer Stelle über Tabellenfunktionen
Neuabtastung.xlsx	Lineare Interpolation und Neuabtastung über VBA-Programm
Polyprox.xlsx	Polynom-Approximation über Matrix-Funktionen MINV
Polyproxa.xlsx	Polynom-Approximation über Matrix-Funktion RGP
Standardfunktionen.xlsx	Anpassung mit den Funktionen in EXCEL (Trendlinie)
Titration.xlsx	Anpassung einer potentiometrischen Titration mit dem SOLVER
Spkgen.xlsx	Spektrengenerator zur Simulation von Spektren