

1. Aufgabe - Voraussetzung

Es ist ein System S gegeben. Um ein mathematisches Modell des Systems zu erstellen, dienen uns als Daten Zahlenpaare, die unser System beim Einsetzen lieferte:

Tabelle 1

x	-2	-1	0	1	2
y	-1,9	-1,1	0	0,9	2,1

Frage

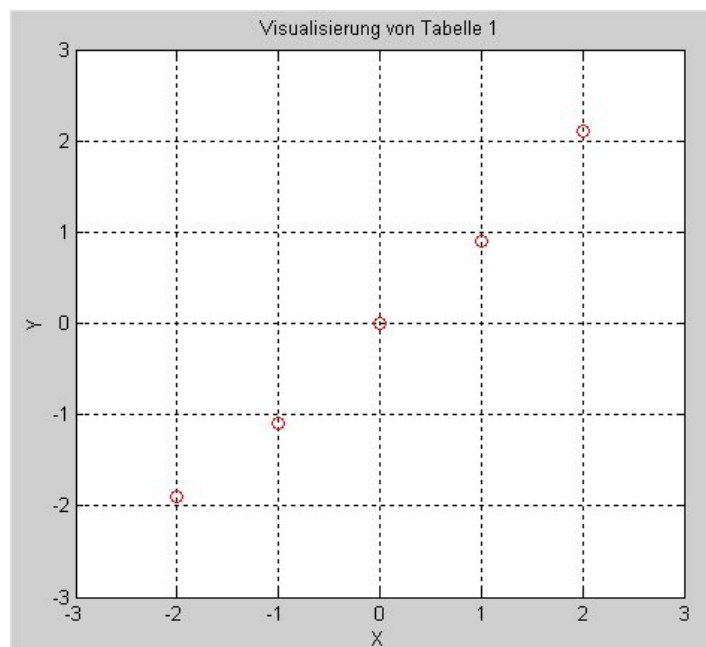
Wie kann dieses System S mathematisch beschrieben werden?

Vorüberlegung

Möchte man ein System mathematisch beschreiben, kann man natürlich beliebig komplexe Formeln verwenden um einen Sachverhalt auszudrücken. Allerdings sollte man dennoch anstreben, mit möglichst wenigen Parametern auszukommen. Die Vorteile davon sind: wenige Daten sind zur Konstruktion nötig, die Formel ist leicht evaluierbar, leicht adaptierbar, feinjustierbar und vor allem verifizierbar. Letzteres ist besonders wichtig um den Grad des Vertrauens festzulegen, mit welcher Sicherheit das Modell das zu beschreibende System abbildet. Zudem sollten die Testpunkte möglichst äquidistant sein. Das wird uns erlauben, den Fehler im Testintervall recht genau abzuschätzen. Genauer zumindest, als zufällige Testpunkte.

Lösungsideen

Da der Mensch talentiert ist bei der optischen Datenverarbeitung, visualisieren wir zunächst unseren Datensatz.



Man stellt fest, dass auf dem Intervall $x = [-2, 2]$ unser System mit einer Toleranz von 0,1 linear verläuft. Es bleibt allerdings offen, wie das System sich außerhalb dieses Intervalls verhalten wird. Damit stellt sich uns die Frage, was unser Anwendungsbereich ist, über welchen Bereich wir eine Aussage zu treffen haben und führt uns zu einer getrennten Betrachtung.

1. Betrachtung des Intervalls [-2, 2]

Da die Testpunkte äquidistant sind, sind wir uns sicher, hier ein Modell zu finden, dass mit hoher Genauigkeit unser System darstellen wird. Die Testpunkte verlaufen nahezu linear. Zur Approximation gibt es mehrere Möglichkeiten.

- Ein Ansatz ist der der Ausgleichsgeraden ($g(x) = ax + b$). Die Minimierung der Fehlerquadrate führt uns zu den Parametern $a = 1$ und $b = 0 \Rightarrow g(x)$ ist die Ursprungsgerade.
- Man könnte auch a und b auch berechnen, indem man über die Steigungsdreiecke mittelt.
- Ein anderer Ansatz ist es, eine Gerade zwischen jeweils zwei der fünf Punkte zu legen, die Fehler zu berechnen und sich dann für diejenige Gerade mit dem geringsten Fehler zu entscheiden. Meistens ist es dabei so, dass eine Gerade durch weiter entfernte Punkte definiert stabiler gegenüber Messfehlern und damit genauer ist.
- Eine weitere Methode: wir „vergessen“ jeweils einen Punkt, bilden aus den übrigen ein beschreibendes Polynom und betrachten dann den Fehler zum weggelassenen Punkt. Wir entscheiden uns dann für das Polynom mit dem geringsten Fehler.
Diese Betrachtung wird uns auch zum Grad führen, den das zu beschreibende Polynom maximal haben sollte.

2. Betrachtung der Intervalle [- , -2] und [2,]

Alle oben beschriebenen Ansätze sind brauchbar, um das System im inneren Intervall zu beschreiben. Bei der Frage allerdings, wie das System sich außerhalb der Testpunkte verhalten wird, besitzen wir keine wirklichen Ansatzpunkte. Es könnte weiterhin linear verlaufen, genauso gut auch periodisch, kubisch oder exponentiell. Auch ist ein Knick oder Sprung nicht auszuschließen und somit gesicherte Aussagen über die Ränder des Systems nicht zu treffen. Es bleibt nur zu diskutieren, welche globale Form das System **wahrscheinlich** haben könnte.

2. Aufgabe - Voraussetzung

Tabelle 2 (Jahr 19XX, Stromverbrauch in MWh)

Jahr	23	25	29	32	44	45	50	60	70	74	78	80
Stromverbrauch der BRD	0,6	1,5	2,0	1,8	5,0	2,5	5,1	12	26	34	40	44

Frage

In dieser Aufgabe ist weniger von Bedeutung, wie man die schon bekannten Datenpunkte durch eine Funktion möglichst genau beschreibt, sondern man möchte vielmehr Zukunftsprognosen machen können, das heißt, Schätzwerte für die nächsten 10 - 20 Jahre erhalten.

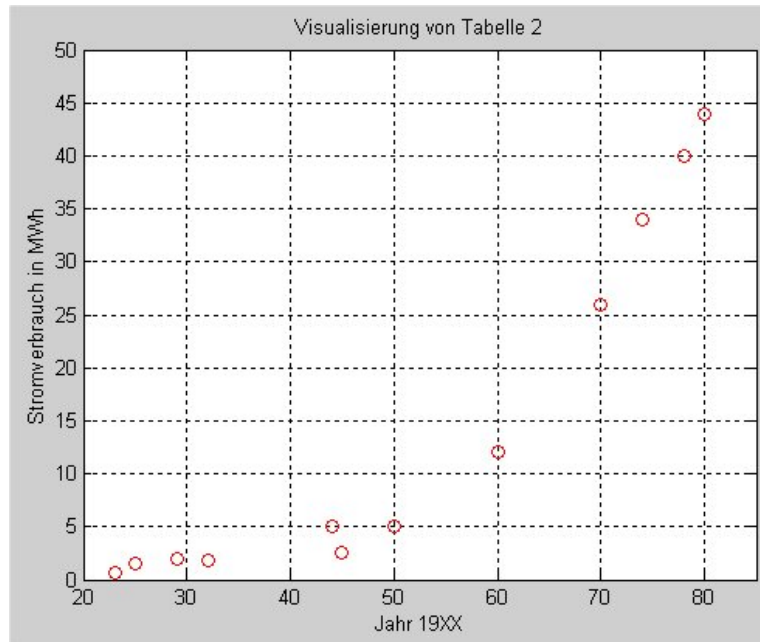
Dies ist von Bedeutung, da zum Beispiel bei Atom-Kraftwerken eine Zeitspanne von ca. 10 Jahren eingeplant werden muss, von der Idee bis zur Inbetriebnahme des Kraftwerkes.

Vorüberlegung

Wiederum stehen wir vor der Frage: lieber geringere Formelkomplexität oder mehr Sicherheit in der Voraussage? Es gilt nach der Wichtigkeit der zu suchenden Werte zu fragen. In diesem Fall sind schlechte Prognosen mit Stromausfall verbunden, eine unangenehme Vorstellung in unserer Zeit. Schon die Funktionsfindung wird in diesem Beispiel zur Frage: mit welchen Daten erzeugen wir welches Modell? D.h. auch: wie sollen wir unsere Daten gewichten? Wie werden wir den Fehler unserer prognostizierenden Funktion abschätzen? Und damit: welche Zuverlässigkeit besitzt unser Modell?

Lösungsideen

Zunächst schaffen wir uns wieder eine optische Anschauung.



- Betrachtet man den Zeitraum ab 1950, so verdoppelt sich ungefähr der Stromverbrauch alle 10 Jahre. Auch eine Annäherung durch alle Punkte deutet auf exponentielles Wachstum hin. Allerdings verzeichnet man bei 1980 einen Rückgang. Würden wir uns also für eine Exponentialfunktion entscheiden, sollten wir eine Sättigung in naher/ferner Zukunft berücksichtigen. Ziehen wir letzteres prinzipiell in Betracht, entfällt zudem ein Polynom als Modellsfunktion.
- Der hintere Datensatz könnte auch linear angenähert werden. Die Methoden dazu siehe erstes Beispiel.
- Komplexeres Modell: logistisches Wachstum ($y' = ay - by^2$). Wird auch in der Realität benutzt. Dieses Modell schließt die Sättigung mit ein (bei $y = a/b$ {= Kapazität} kein Wachstum mehr).

Hat man ein Modell gefunden, muss man noch seine Prognosefähigkeit verifizieren. Abgesehen von der schon oben beschriebenen Methode, jeweils einen Punkt zu vergessen, aus den restlichen eine Funktion zu erstellen und dann den entstandenen Fehler zu betrachten, könnte man alle Werte ab einem gewissen Zeitpunkt vergessen. Man versucht z.B. ein Modell bis 1970 zu entwickeln und schaut, welcher Fehler dabei entstünde, wenn man nun anhand dieses Modells Aussagen über 1980 gemacht hätte.

Diese Methodik bleibt allerdings wie jede andere auch auf langfristige Prognosen (> ca. 15 Jahre) ungewiss, auf noch längere Sicht riskant und je nach Bedeutung sogar gefährlich. Zu viele andere Faktoren spielen noch eine Rolle, die theoretisch mit in die Modellbetrachtung einbezogen werden müssten, wie

- Bevölkerungswachstum
- Wirtschafts- und Industriewachstum
- Produktionswachstum von stromverbrauchenden Geräten
- ...

Faktoren, die zwar längerfristige Voraussagen eher verkomplizieren, allerdings zur Modellfindung für die nahe Zukunft sogar hilfreich sind.

Bleibt noch zum Schluss zu sagen, dass wohl doch immer ein Schuss gute Intuition dazugehört.