

Zeitreihen

Was sind Zeitreihen?

Unter einer Zeitreihe versteht man die Beschreibung des Zustands eines Systems in Abhängigkeit zur Zeit t . Der Zustand wird mithilfe einer Zustandsvariablen (K) ausgedrückt, welche beispielsweise einen Kontostand beschreiben kann. Das System besitzt einen Ausgangszustand $K(t_0)$. Der Zustand kann nicht kontinuierlich angegeben werden, sondern nur zu diskreten Zeitpunkten $t_0, t_1, t_2, \dots, t_i$. Diese Zeitpunkte sind äquidistant, also immer um den selben Zeitschritt voneinander entfernt ($t_i + h = t_{i+1}$). Das System enthält einen oder mehrere Systemparameter p , die die Eigenschaften des Systems beschreiben und die in Abhängigkeit von der Zeit stehen können. $K(t_0)$ und p sind durch Umgebungsbedingungen gegeben oder gegebenenfalls frei wählbar.

Der Zustand zu einem gegebenen Zeitpunkt wird gewöhnlicherweise mit der Formel

$$K(t_{i+1}) = K(t_i) + \Delta K(t_i)$$

errechnet, wobei $\Delta K(t_i)$ die Änderungsrate zum Zeitpunkt t_i angibt, das heißt den Wert, um den sich der Zustand nach dem Zeitschritt i ändert. Je nach Wahl der Funktion für $\Delta K(t_i)$ können hier viele verschiedene Systeme abgebildet werden.

$\Delta K(t_i)$ könnte beispielsweise wie folgt aussehen: $\Delta K(t_i) = p \cdot K(t_i)$, wodurch ein einfaches Zinswachstum mit dem Zinssatz p ausgedrückt wird.

Beispiele

Die Abbildungen sind auf den Seiten 3 bis 5 zu finden.

Konto mit Zinsen (Abbildung 1)

Ein einfaches Beispiel für Zeitreihen ist die Kapitalentwicklung eines Bankkontos mit einem festen Zinssatz.

Hier entspricht der Zustand K dem aktuellen Kontostand, $K(t)$ dem Kontostand zum Zeitpunkt t . $K(t_0)$ ist gegeben als der Kontostand zu Beginn der Zeitreihe, also ein bestimmtes Kapital, das einmalig eingezahlt wird.

Die Funktion $\Delta K(t_i)$ lautet wie folgt:

$$\Delta K(t_i) = p \cdot K(t_i)$$

Der Systemparameter p gibt den Zinssatz an.

Dieses Beispiel ist gut als Zeitreihe auszudrücken, weil Zinsen von Banken gewöhnlich in jährlichen Abständen, also zu diskreten, äquidistanten Zeitpunkten ausgezahlt werden, und weil die Kapitalentwicklung durch Zinsen den Gesetzen des Wachstums folgt.

In der praktischen Umsetzung des Modells hat sich gezeigt, dass die Kurve einer exponentiellen Wachstumskurve gleicht.

Entwicklung einer Epidemie (Abbildung 2)

Epidemien entwickeln sich – mathematisch gesehen – nach dem folgenden Prinzip: Eine gewisse Anzahl an Individuen ist infiziert, jeder von ihnen trifft pro Tag eine Anzahl anderer Individuen und steckt sie an.

Nach einer gewissen Zeit jedoch sind schon sehr viele Individuen infiziert, und man kann davon ausgehen, dass jedes von ihnen auch bereits infizierte Individuen trifft. Somit infiziert es nicht alle Individuen, die es trifft, sondern nur einen bestimmten Prozentsatz.

Festzulegen sei hier, dass $K(t_0)$ dem Wert der anfangs infizierten Personen entspricht, und p den Anteil an Individuen von der Gesamtzahl, die jedes einzelne Individuum pro Zeitschritt ansteckt. Hierbei treten die Regeln des logistischen Wachstums in Kraft. Dies wird in der Funktion $\Delta K(t_i)$ ausgedrückt:

$$\Delta K(t_i) = p \cdot K(t_i) \cdot (L - K(t_i))$$

L entspricht der Gesamtzahl aller Individuen (also sowohl infizierte als auch gesunde). Somit entspricht $L - K(t_i)$ der Anzahl gesunder Individuen. Also wird der Prozentsatz p zunächst mit der Anzahl infizierter Personen multipliziert, und danach mit der Anzahl gesunder Individuen, was dazu führt, dass die Wachstumsrate geringer wird, je mehr Individuen schon infiziert sind.

Man beachte, dass p ein sehr kleiner Wert ist. Beispielhaft soll nun angenommen werden, dass die Population 100 000 Individuen beträgt und jedes infizierte Individuum zwei weitere ansteckt.

Hierbei ist $p = \frac{2}{100\,000}$, also 0,00002.

Entwicklung der Weltbevölkerung (Abbildung 3)

Ein ähnliches Modell wird verwendet, um die Entwicklung der Weltbevölkerung zu beschreiben oder vorauszusagen. $K(t_0)$ ist die Bevölkerung zu Beginn der Zeitreihe, p der Prozentsatz, um den die Bevölkerung pro Zeitschritt im Durchschnitt wächst, und L die Anzahl Menschen, die schätzungsweise auf der Erde leben können, ohne ihr eigenes Überleben gegenseitig zu behindern.

Es gilt wieder $\Delta K(t_i) = p \cdot K(t_i) \cdot (L - K(t_i))$.

Da die Entwicklung der Weltbevölkerung (Geburten- und Sterberate) prozentual nahezu konstant ist, kann p konstant gewählt werden. Zu beachten ist hier, dass L sich verändern kann. Wie schon beschrieben, gibt es die Anzahl Menschen an, die auf der Erde leben können, ohne sich gegenseitig am Überleben zu hindern. Dies schließt z. B. mit ein, dass genügend Lebensmittel für jeden zur Verfügung stehen, oder der Wohnkomfort im Mittel so hoch ist, dass die verfügbare Wohnfläche auf der Erde ausreicht. Man halte sich nun beispielsweise die Entdeckung Amerikas vor Augen, welches vorher sehr dünn besiedelt war, wohingegen Europa einen relativ dichten Besiedelungsquotienten hatte. Nun entdeckten die Europäer Amerika, was bedeutete, dass sie wieder eine sehr große Menge an Land zum Bewohnen und Bestellen zur Verfügung hatten. Somit hat sich im Modell L erhöht.

Im Schaubild kann man hier einen Knick in der Kurve erkennen.

Im Modell wurde für $\Delta K(t_i)$ die Formel

$$\Delta K(t_i) = p \cdot K(t_i) \cdot \frac{(L - K(t_i))}{L}$$

gewählt, um im letzten Faktor nicht eine absolute Zahl, sondern einen Anteil zu erhalten. Dies dient nur Demonstrationszwecken, da ansonsten die Zahlen zu groß und der Knick zu undeutlich würden.

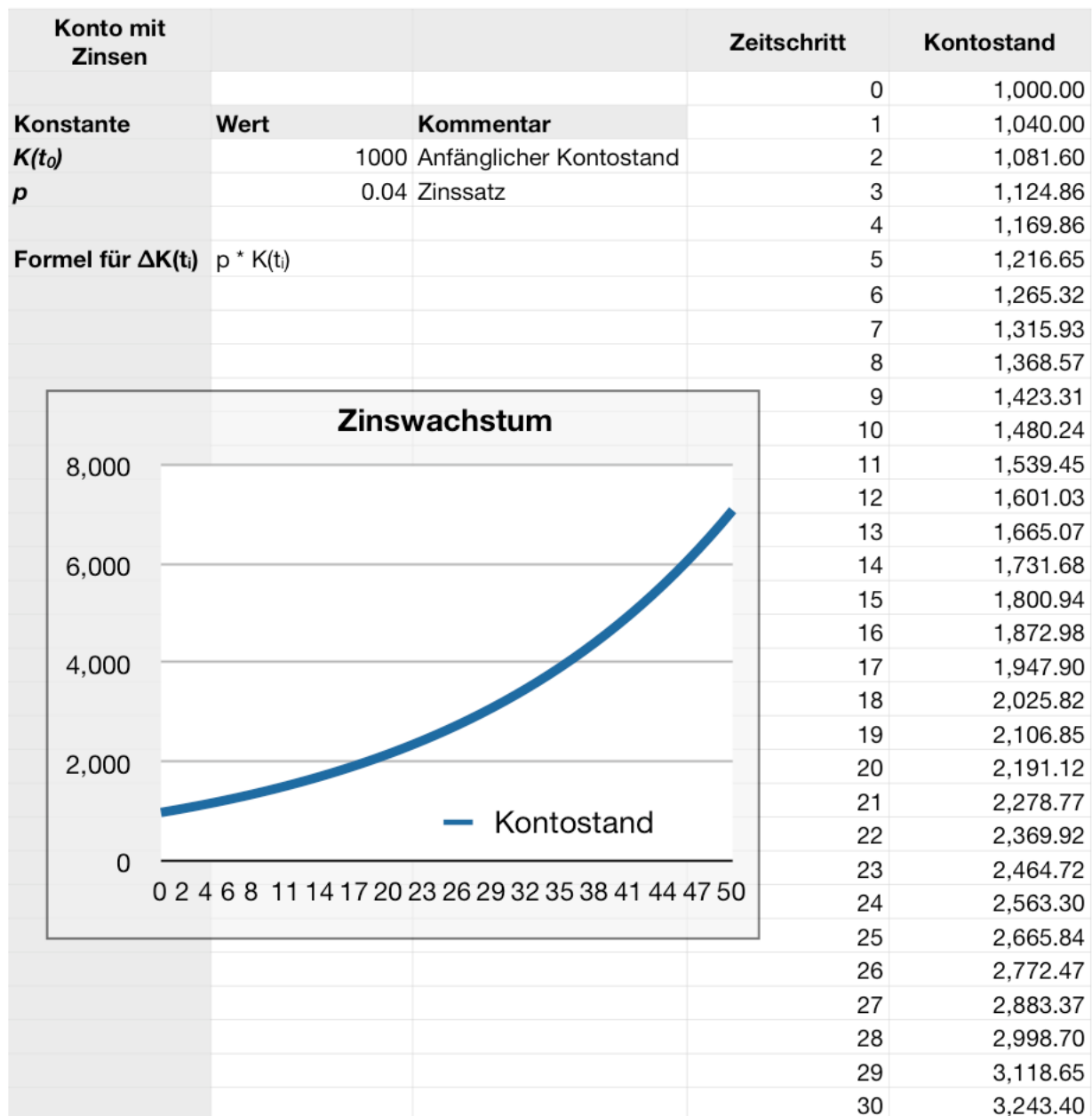


Abbildung 1: Konto mit Zinsen

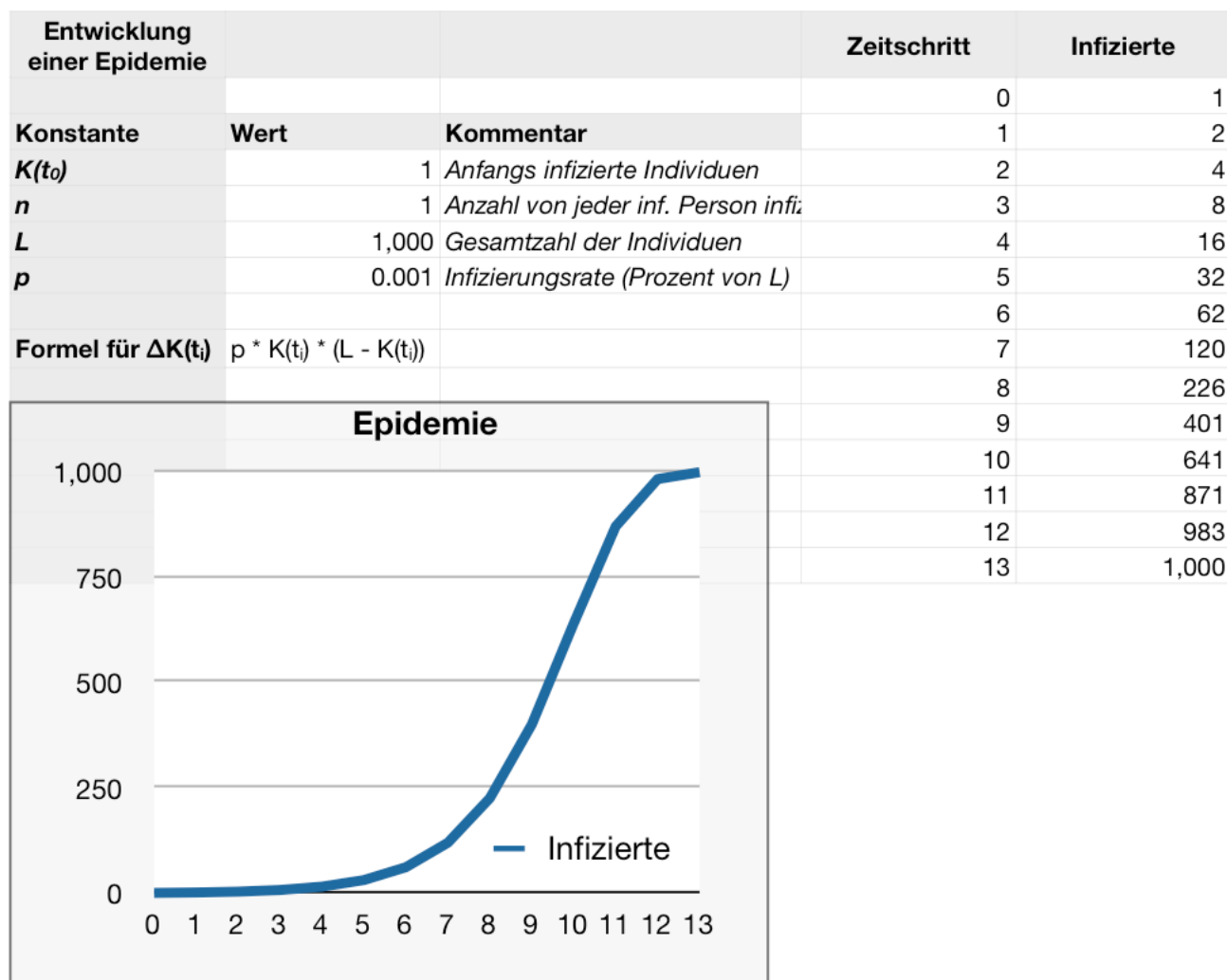


Abbildung 2: Entwicklung einer Epidemie

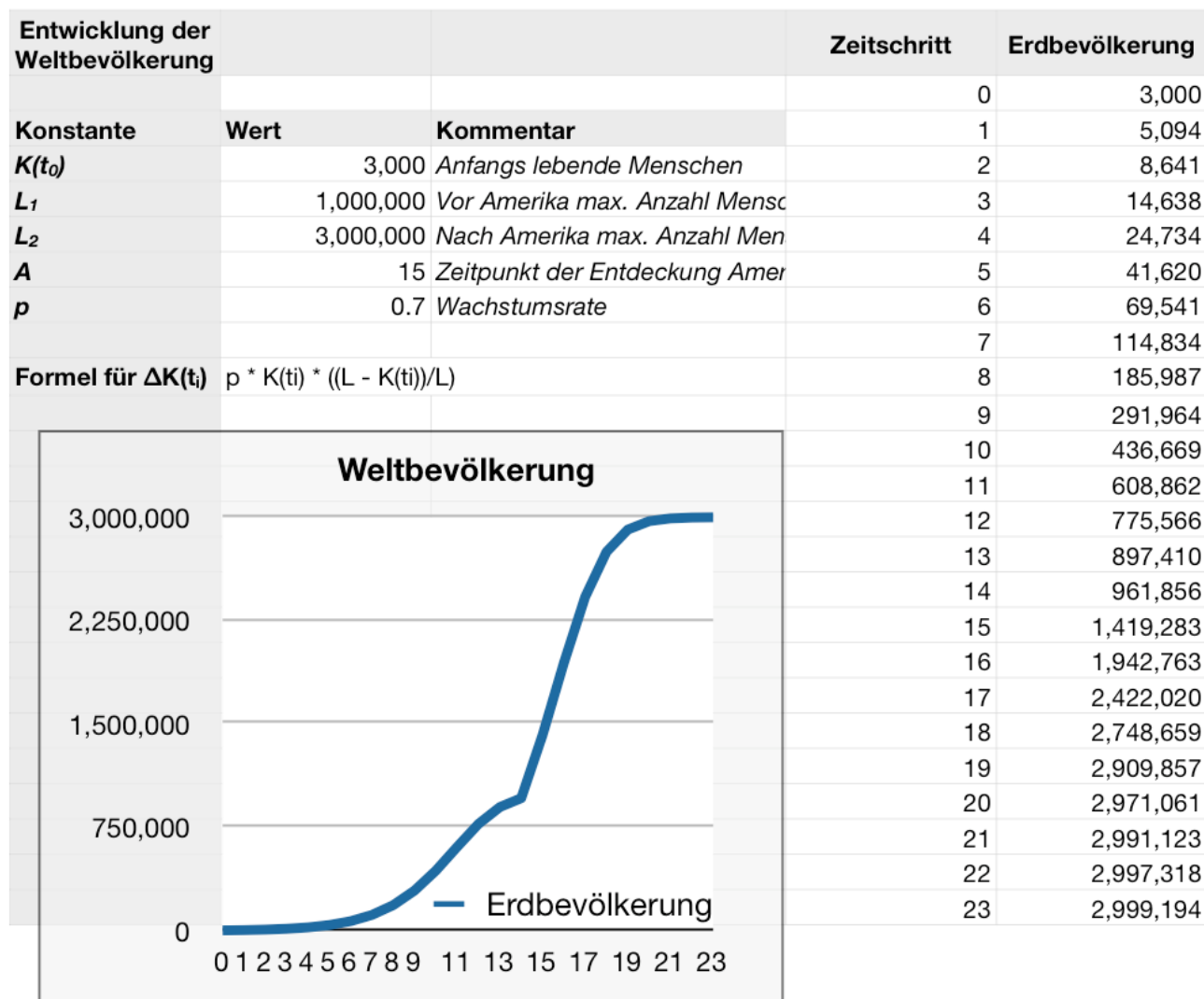


Abbildung 3: Beispielhafte Berechnung der Entwicklung der Weltbevölkerung

Programmiertechnische Umsetzung

Ansatz und Beschreibung

Der Grundansatz ist es, eine Klasse *TimeSeries* für die generelle Berechnung von Zeitreihen anzulegen. In ihr werden alle Werte für K als Array abgelegt, sie kümmert sich auch um die Speicher-verwaltung hierfür.

Diese Klasse berechnet allerdings nicht ΔK – hierfür wurde eine eigene Klasse *Delta* mit einer virtuellen Methode zum Berechnen von ΔK anhand von t und K angelegt. *Delta* selbst berechnet jedoch nichts, hierfür muss eine Klasse erstellt werden, die von *Delta* erbt, und in der die vormals virtuelle Methode ΔK berechnet und zurückgibt.

Da dieses Problem durch Klassenvererbung angegangen wurde, kann der Programmierer der Zeitreihe selbst Parameter geben, indem er sie einfach in der von *Delta* erbenden Klasse definiert.

Die Klasse *TimeSeries* berechnet nun anhand eines übergebenen Startwertes und *Delta* $\Delta K(t_0)$, addiert gemäß der oben aufgestellten Formel

$$K(t_{i+1}) = K(t_i) + \Delta K(t_i)$$

$K(t_0)$ und speichert das Ergebnis als K_{t_i} , und so fort.

In beiden Klassen wurden C++-Templates verwendet, was bedeutet, dass die Klassen für sämtliche Datentypen verwendbar sind, also auch eigene Klassen, die dynamische Systeme oder Ähnliches beschreiben. Sie müssen lediglich den Operator „+“ überladen oder unterstützen.

Anwendung

Zunächst wird eine von der Klasse *Delta* abgeleitete Klasse geschrieben. Diese muss die Methode *calculate* enthalten, deren Deklaration wie folgt aussieht:

```
virtual Object calculate (int t, Object *stock)
```

wobei *Object* ein Klassentemplate ist. *stock* entspricht K .

In dieser Methode wird ΔK berechnet und zurückgegeben.

Nun erstellt man ein neues *TimeSeries*-Objekt, dem als Templateargument der gewünschte Datentyp (*int*, *double*, eigene Klassen, etc.) übergeben wird. Als Argumente nimmt der Konstruktor einen Startwert (K_{t_0}) und einen Zeiger auf die neu erstellte *Delta*-Klasse entgegen.

`TimeSeries::next()` berechnet den K -Wert für den nächsten Zeitschritt und gibt ihn zurück.

`TimeSeries::calculateSteps(int n)` berechnet die nächsten n Schritte und gibt die Werte per `TimeSeries::printStep(int t, Object result)` aus. Dies ist eine virtuelle Funktion und kann so durch Vererbung ersetzt werden, um die Ausgabe abzuändern oder die Daten anderweitig zu verändern.