

Florian Heinrichs

Algorithmen, Zufall, Unsicherheit – und Pizza!

Wie Mathematik uns hilft,
alltägliche Entscheidungen
zu treffen

SACHBUCH



Springer

Algorithmen, Zufall, Unsicherheit – und Pizza!

Florian Heinrichs

Algorithmen, Zufall, Unsicherheit – und Pizza!

Wie Mathematik uns hilft,
alltägliche Entscheidungen
zu treffen

 Springer

Florian Heinrichs
Professor für Data Science und Statistik Fachbereich
Medizintechnik und Technomathematik
FH Aachen – University of Applied Sciences
Jülich, Deutschland

ISBN 978-3-662-69171-7 ISBN 978-3-662-69172-4 (eBook)
<https://doi.org/10.1007/978-3-662-69172-4>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://portal.dnb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2024

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Iris Ruhmann

Springer ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Wenn Sie dieses Produkt entsorgen, geben Sie das Papier bitte zum Recycling.

Für Klaus Heinrichs

Dank

Dieses Buch wäre nicht entstanden ohne die Hilfe und Unterstützung unzähliger Personen. Meine alltäglichen Erfahrungen mit der Mathematik sind die Basis dieses Buches. Die Erfahrungen sind jedoch untrennbar verknüpft mit Menschen – insbesondere meinen Eltern, Brüdern, Freunden und Lehrern. Es ist unmöglich, all diejenigen zu erwähnen, die einen impliziten Einfluss auf dieses Buch hatten. Trotzdem möchte ich es versuchen und mich speziell bei denjenigen bedanken, deren Unterstützung in den letzten Jahren dieses Buch maßgeblich beeinflusst hat.

Mein Dank gilt Holger Dette, meinem Doktorvater, der mich vom ersten Semester meines Studiums bis zum Ende meiner Promotion begleitet hat. Von ihm durfte ich vieles lernen, das ich heute über die Mathematik, Zufall und Wahrscheinlichkeiten weiß. Durch ihn bin ich zur Zeitreihenanalyse gekommen, die meinen Blick auf die Welt geprägt hat.

Weiter möchte ich mich bei Axel Bücher bedanken, mit dem ich zum Ende meines Studiums und während der Promotion eng zusammenarbeiten durfte. Ich konnte fachlich und menschlich viel von ihm lernen – von statistischen Me-

VIII Dank

thoden bis zu speziellen Klettertechniken. Seine Konsequenz und Prinzipientreue haben mich immer beeindruckt.

Besonders geprägt haben mich auch der fachliche (und weniger fachliche) Austausch mit Josua Gösmann, Kevin Klinkenborg, Rafael Kurek, Daniel Meißner und Martin Dunsche. Viele Erkenntnisse aus unseren Gesprächen an der Universität und beim Sport sind in dieses Buch eingeflossen.

Die Fertigstellung des Buches war ein langer Prozess. Camilla Szymanski hat mich dabei immer wieder angespornt und motiviert, weiter zu schreiben und das Ziel nicht aus den Augen zu verlieren.

Außerdem danke ich Iris Ruhmann, die dem Konzept dieses Buches von Anfang an offen gegenüberstand und mit ihrem professionellen und hilfreichen Feedback geholfen hat, aus einem unfertigen Manuskript ein (hoffentlich) unterhaltsames Buch zu machen. Mein Dank gilt auch Betina Saglio, die das Buch als Projektmanagerin betreut hat.

Als Letztes möchte ich mich besonders bei meiner Frau Stephanie bedanken, die mich von der ersten Idee bis zur Veröffentlichung des Buches unterstützt und mir die Freiräume geschaffen hat, die es mir erlaubten, an diesem und anderen Projekten zu arbeiten. Sie hat die ersten Entwürfe des Manuskriptes gelesen und durch ihr Feedback maßgeblich verbessert. Ohne ihr unendliches Verständnis und ihre kontinuierliche Unterstützung wäre die Erstellung dieses Buches kaum möglich gewesen. Vielen Dank!

Inhaltsverzeichnis

Einleitung	1
Algorithmen: Die Grundidee der systematischen Herangehensweise	3
Ordnung in der Plattensammlung – Sortieralgorithmen	8
Schnelle Suche für schnelle Forschung – Suchalgorithmen	16
It's a (perfect) Match – Stabile Matchings	21
Optimierungsalgorithmen: Fundierte Entscheidungen treffen	29
Die kürzeste Strecke finden – Wegoptimierung	31
Eine Wanderung durch die Berge – Optimierung stetiger Funktionen	38
Zufall und Wahrscheinlichkeiten: Unsicherheit modellieren	47
Glücksspiele und Casinos – Wahrscheinlichkeiten ...	48
Quantifizierung des Zufalls – Erwartungswerte	57
Risikobewertung und zufällige Schwankungen – Varianz	64

Wie viele Getränke brauchen wir? – Schätzer und Konfidenzbereiche	69
Welche Mannschaft ist besser? – Hypothesentests . . .	77
Wann ist das Benzin günstig? – Regression und Zeitreihenanalyse	86
Aberglaube und Warenkörbe – Korrelation und Kausalität	92
Verspätung und extreme Ereignisse – Ausreißerdetektion	100
Veränderungen und Flugverkehr – Strukturbruchanalyse	108
Algorithmen und Zufall: Optimale Entscheidungen treffen trotz unsicherer Umstände	115
Verhandlungen und Parkplatzsuche – Optimal Stopping	121
Sicherheit und Fußballfans – Klassifikatoren	130
Der Fluch der Dimensionalität – Dimensionsreduktion	139
Sitzpläne und Sightseeing – Clustering	148
Suche nach dem besten Film – Recommender Systems	155
Pizzerien und Entscheidungen – Reinforcement Learning	158
Trainingspläne und Filmvorschläge – Q-Learning . . .	164
Vorlesungen und Rundreisen – Evolutionäre Algorithmen	170
Epilog: Ein Tag ohne Mathematik	177
Glossar	179
Literatur	187



Einleitung

In meiner Lieblingspizzeria kostet eine große Pizza Margherita 6 €, bei einem Durchmesser von 30 cm. Wer besonders viel Hunger hat, kann auf „Angebot 3“ zurückgreifen: ein Pizzablech der Größe 40 cm × 60 cm für 24 €.

In der Schule lernen wir, die Flächen von Vierecken und Kreisen zu berechnen, und können so herausfinden, welche Option günstiger ist: ein Pizzablech oder vier Pizzen? Tatsächlich kosten vier Pizzen genauso viel wie ein Pizzablech, haben jedoch die größere Fläche. Daher ist das „Angebot“ letztlich teurer als die einzelnen Pizzen.

Offensichtlich ist Schulmathematik im Alltag nützlich, doch sie wird oft schlecht motiviert. Statt die Größe von Pizzen zu vergleichen, geht es in der Schule häufig um Fragen wie „Wie groß ist eine Wiese, die 4 m lang und 6 m breit ist?“ oder „Welche Fläche kann mit einem 20 m langen Zaun höchstens abgesperrt werden?“.

Durch irrelevante Beispiele entsteht eine Lücke zwischen Mathematik und Alltag, durch die viele Menschen ihr Interesse an der Mathematik verlieren. Doch die Bedeutung der

Mathematik nimmt immer weiter zu. Sie steckt in Technologien, die wir täglich nutzen. Wenn Unternehmen wie Google oder Amazon neue Rekordumsätze erwirtschaften, heißt es in den Nachrichten oft, Daten seien das neue Gold. Doch Daten sind für Tech-Konzerne nur deswegen wertvoll, weil sie mit Hilfe mathematischer Methoden effizient ausgewertet werden können.

Um Technologien und Entwicklungen verstehen zu können, die unsere moderne Gesellschaft beeinflussen, benötigen wir ein intuitives Verständnis einiger grundlegender mathematischer Ideen. In diesem Buch geht es um diese Grundlagen, vor allem aber um ihre praktische Anwendung und Konzepte, die sich im Alltag nutzen lassen.

So kann die Mathematik nicht nur genutzt werden, um die Flächen verschiedener Pizzen zu berechnen, sondern auch dabei helfen, viel relevantere Fragen zu beantworten:

- Bei welcher Pizzeria sollten wir bestellen? Und welche Pizza?
- Ab welcher Entfernung von unserem Ziel sollten wir nach einem Parkplatz suchen? Ab wann sollten wir auf dem nächsten freien Platz parken?
- Was ist der kürzeste Weg zum neuen Arbeitsplatz?
- Wie viele Getränke brauchen wir für eine Party?

Das Ziel dieses Buches ist aufzuzeigen, dass ein Alltag ohne Mathematik kaum vorstellbar ist, wo sie uns begegnet, wo wir (oft unbewusst) täglich mathematische Ideen benutzen und wie wir mit Hilfe von Mathematik bessere Entscheidungen treffen können.



Algorithmen: Die Grundidee der systematischen Herangehensweise

Einer der ersten – und nach wie vor besten – Filme, die ich im Kino gesehen habe, ist *Star Wars: Episode I – Die dunkle Bedrohung*. Schon als Kind fand ich die aufwendigen Kampfszenen mit ihren Spezialeffekten beeindruckend. Besonders faszinierte mich die *Droidenarmee*, eine Armee von Kampfrobotern, die glücklicherweise kurz vor einer verheerenden Schlacht gestoppt werden konnte. Doch auch in anderen dystopischen Science-Fiction-Filmen tauchen immer wieder Roboter auf, die erst die Menschheit versklaven (Terminator) und sie dann wieder befreien wollen (Terminator 2), wobei diese Roboter von Algorithmen gesteuert werden.

Inspiziert von Hollywood-Blockbustern entstehen Nachrichtenartikel und wissenschaftliche Arbeiten mit Titeln wie *Angriff der Algorithmen* [1], *Algorithmen: Wie und warum sie Menschen diskriminieren*, [2] oder *Diskriminierungsrisiken von Algorithmen* [3].

Diese Titel suggerieren, dass Algorithmen eine Form von Bewusstsein haben, Entscheidungen treffen, danach

handeln und letztlich auch verantwortlich für ihr Handeln gemacht werden können. In den letzten Jahren hat sich der Begriff *Algorithmus* zu einem Schlagwort entwickelt, das in allen möglichen Situationen genutzt wird – insbesondere dann, wenn Entscheidungsträger aus Politik und Wirtschaft die Verantwortung für Fehler und menschliches Versagen auf *den Algorithmus* abwälzen wollen. Doch was steckt hinter diesem Schlagwort? Was sind Algorithmen? Stehen wir kurz vor der Unterwerfung der Menschheit durch menschenähnliche Roboter?

Ein Algorithmus ist eine endliche Folge von Schritten zum Lösen eines Problems [4]. Die *Schritte* des Algorithmus beschreiben präzise Handlungsanweisungen und sind bei jeder Ausführung des Algorithmus dieselben. Dabei ist es wichtig, dass der Algorithmus nur aus endlich vielen Schritten besteht, damit er irgendwann endet bzw. *terminiert* und nicht für immer weiterläuft.

Mit dieser allgemeinen Definition ist es leicht zu sehen, dass uns Algorithmen auch außerhalb von Computern und Smartphones häufig im Alltag begegnen. Kochrezepte, Wegbeschreibungen und Bedienungsanleitungen sind klassische Beispiele für Algorithmen. Wenn ein Fahrschüler die unterschiedlichen Schritte erlernt, um ein Auto zu starten und in Bewegung zu setzen, erlernt er einen Algorithmus. In der ersten Zeit wird er jeden dieser Schritte bewusst durchführen, bis der Algorithmus irgendwann in sein Unterbewusstsein übergegangen ist. Bauanleitungen für Möbelstücke sind ebenso Algorithmen wie Pflegeanleitungen für Pflanzen.

Durch die Beispiele wird klar, weshalb Algorithmen terminieren sollten. Ein Kochrezept, das aus unendlich vielen Schritten besteht, ist ebenso wenig praktisch wie eine Wegbeschreibung, mit der wir das Ziel niemals erreichen. Außerdem erwarten wir ein bestimmtes Resultat. Wenn wir

beispielsweise zum Backen ein Rezept für einen Streuselkuchen nutzen und bei einem Mal einen Apfelkuchen und beim anderen Mal einen Schokoladenkuchen bekommen, wäre das Rezept nicht dafür geeignet, unser Ziel zu erreichen, einen Streuselkuchen zu backen, weil die Handlungsanweisungen nicht präzise genug sind.

Algorithmen sind nicht neu. So waren Algorithmen, wie das Sieb des Eratosthenes oder der euklidische Algorithmus, schon im antiken Griechenland bekannt – wenngleich nicht unter diesen Bezeichnungen. Im 19. und 20. Jahrhundert formalisierten Mathematiker den bis dahin nur informell definierten Algorithmus basierend auf einer theoretischen Konstruktion, der Turing-Maschine, die maßgeblich zur Entwicklung von modernen Computern beitrug [5].

Häufig werden Algorithmen im *Pseudocode* dargestellt. Pseudocode ist eine Mischung aus natürlicher Sprache und Programmcode, der durch seine einfache Struktur gut dazu geeignet ist, auch komplizierte Algorithmen übersichtlich darzustellen. Algorithmen 1 und 2 sind jeweils Beispiele für einen terminierenden und einen nichtterminierenden Algorithmus in Pseudocode. Im ersten Schritt initialisiert Algorithmus 1 die *Variable* x mit dem Wert 0 („setze x gleich 0“). Die nächsten beiden Schritte bilden eine sogenannte *Schleife*. Solange die Bedingung aus Schritt 2 erfüllt ist, wird der Befehl in Schritt 3 ausgeführt. In diesem Fall wird also x schrittweise um 1 erhöht, bis es einen bestimmten Wert annimmt – in diesem Beispiel den Wert 10. Die Bedingung der Schleife ist durch die Struktur „Solange ..., wiederhole:“ gekennzeichnet und der Hauptteil (in Zeile 3) dadurch, dass er eingerückt ist. Algorithmus 2 hat eine ähnliche Struktur, jedoch wird hier der Schritt in Zeile 3 unendlich oft wiederholt, da in Zeile 2 keine Abbruchbedingung definiert wird, sodass der Algorithmus nicht endet.

Algorithmus 1 Beispiel für einen Algorithmus in Pseudocode. Der Algorithmus beginnt bei 0 und zählt bis 10. Dabei wird der Wert von x in jedem Schritt um 1 erhöht

1. Setze x gleich 0
2. Solange x kleiner oder gleich 10 ist, wiederhole:
3. Erhöhe den Wert von x um 1

Algorithmus 2 Beispiel für einen Algorithmus, der nicht terminiert. Wie Algorithmus 1 beginnt er bei 0 und erhöht x in jedem Schritt um 1 – ohne jemals zu enden

1. Setze x gleich 0
2. Wiederhole:
3. Erhöhe den Wert von x um 1

Neben den Handlungsanweisungen haben Algorithmen häufig eine Ein- und Ausgabe, also Komponenten, die dem Algorithmus vor Beginn mitgeteilt werden müssen (Eingabe), und Ergebnisse, die aus dem Algorithmus resultieren (Ausgabe). Mit Hilfe eines Kochrezepts (Algorithmus), können Sie aus Zutaten (Eingabe) eine Speise (Ausgabe) zubereiten. Ein alltagsnahes, nützliches Beispiel für Pseudocode ist ein Rezept für Crêpes (wie in Algorithmus 3). Der Hauptteil der Schleife in Algorithmus 3 besteht aus drei Schritten, die wiederholt werden, bis der Teig leer ist.

Algorithmus 3 Beispiel für ein Backrezept zur Zubereitung von Crêpes im Pseudocode. Schritte 1 und 2 werden nur einmal durchgeführt, wohingegen die Schritte 4 bis 6 durchgeführt werden, solange die Bedingung in Schritt 3 erfüllt ist

Eingabe: 2 Eier, 250 g Mehl, 500 ml Milch, 1 Prise Salz

Ausgabe: Crêpes

1. Mische die Zutaten zu einem Teig
2. Erhitze eine Pfanne auf mittlerer Stufe
3. Solange Teig vorhanden, wiederhole:
4. Gib so viel Teig in die Pfanne, bis der Boden bedeckt ist
5. Wende den Crêpe, sobald der Teig ausreichend fest ist
6. Serviere den fertigen Crêpe

Offensichtlich treffen Algorithmen nicht selbständig Entscheidungen oder wählen bewusst aus unterschiedlichen Handlungsalternativen. Daher ist es mehr als fragwürdig, wenn die Verantwortung für negative Konsequenzen, wie Diskriminierung oder Manipulation von Menschen, die aus dem Einsatz von Algorithmen resultieren, auf ebendiese abgewälzt wird. Die Verantwortung liegt bei den Entwicklern und Anwendern von Algorithmen. Natürlich ist dies dem Autor der eingangs erwähnten Studie, die die Antidiskriminierungsstelle des Bundes mit „Diskriminierungsrisiken von Algorithmen“ bewirbt, bekannt. Die Studie trägt den Titel „Diskriminierungsrisiken durch Verwendung von Algorithmen“ [6]. Durch das Weglassen der Worte „durch Verwendung“ bei der Veröffentlichung der Studie hat die Antidiskriminierungsstelle die Aussage des Titels verändert.

In diesem Kapitel werden wir uns mit wichtigen Algorithmen beschäftigen, die uns im Alltag begegnen – vom Sortieren der Plattensammlung über die Suche eines Buches bis zum Finden der optimalen Arbeitsstelle.

Das Wichtigste in Kürze

- Ein Algorithmus ist eine endliche Folge von Schritten zum Lösen eines Problems.
- Algorithmen treffen keine bewussten Entscheidungen – die Verantwortung liegt bei den Entwicklern und Anwendern der Algorithmen.
- Beispiele für Algorithmen sind Kochrezepte, Wegbeschreibungen und Bedienungsanleitungen.

Ordnung in der Plattensammlung – Sortieralgorithmen

Seit einigen Jahren erfreuen sich Boxen mit der Aufschrift „zu verschenken“ in meiner Nachbarschaft immer größerer Beliebtheit. Die Idee ist simpel: Dinge, die nicht mehr gebraucht werden, die allerdings zu gut zum Wegwerfen sind, können in die Kiste gelegt werden und so einen neuen Besitzer finden, der eine Verwendung für die Gegenstände hat. Bei einem Spaziergang fand ich den originalverpackten Soundtrack von Django Unchained als Doppelalbum in einer dieser Boxen und konnte als leidenschaftlicher Plattensammler der Versuchung nicht widerstehen. Als ich die Schallplatte in meine Sammlung einsortieren wollte, musste ich feststellen, dass die restlichen Platten durch mehrere Umzüge völlig durcheinander gemischt waren. Es war also an der Zeit, die gesamte Plattensammlung zu sortieren.

Ähnliche Probleme wie das Sortieren einer Plattensammlung begegnen uns beim Sortieren von Büchern, Kleidung oder Werkzeugen. So unterschiedlich wie die Anwendungsfelder können auch die Kriterien sein, nach denen wir sortieren. Schraubenschlüssel oder Schuhe können Sie beispielsweise nach ihrer Größe sortieren, Gewürze im Regal

nach der Häufigkeit ihrer Nutzung und Schallplatten nach Interpret. Selbstverständlich können Sie Gegenstände anhand unterschiedlicher Kriterien sortieren. Eine Büchersammlung könnten Sie zum Beispiel alphabetisch nach Autor, thematisch oder farblich sortieren. Wichtig ist nur, dass die Ordnung eindeutig ist – so muss zum Beispiel die Reihenfolge der Farben vor dem Sortieren klar sein. Doch wie können Sie in all diesen Szenarien möglichst effizient Ordnung schaffen? Ein Ansatz zum effizienten Sortieren sind sogenannte *Sortieralgorithmen*. Damit diese Algorithmen unabhängig von dem konkreten Anwendungsfall benutzt werden können, machen sie Gebrauch von dem abstrakten Begriff einer *Liste*, die sortiert werden soll. Je nach Anwendung kann die zu sortierende Liste beispielsweise eine Menge von Büchern oder eine Plattensammlung sein.

Ein intuitiver Sortieralgorithmus ist der *Bubblesort*-Algorithmus, dessen Pseudocode Sie in Algorithmus 4 finden [7]. Wenn Sie beispielsweise Schallplatten von Die Toten Hosen (DTH), Nirvana (NIR), Manu Chao (MAN) und Django Unchained (DJA), kurz DTH – NIR – MAN – DJA, alphabetisch sortieren möchten, können Sie die Schallplatten jeweils von links nach rechts vergleichen und die Ordnung von benachbarten Platten korrigieren – wie in Abb. 1:

- a) DTH kommt vor NIR und die Platten bleiben in dieser Reihenfolge:
DTH – NIR – MAN – DJA
- b) NIR kommt nach MAN und die beiden Platten werden getauscht:
DTH – MAN – NIR – DJA (1. Wechsel in Abb. 1)
- c) NIR kommt nach DJA und die beiden Platten werden getauscht:
DTH – MAN – DJA – NIR (2. Wechsel in Abb. 1)

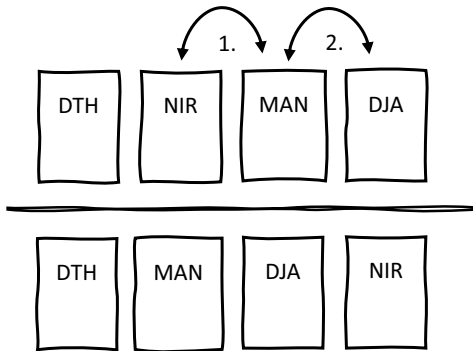


Abb. 1 Alphabetisches Sortieren einer Plattensammlung – Vergleich von je zwei benachbarten Schallplatten

Algorithmus 4 Pseudocode vom Bubblesort-Algorithmus zum Sortieren einer Liste mit n Objekten

Eingabe: Unsortierte Liste mit n Objekten $X_1, X_2, X_3, \dots, X_n$

Ausgabe: Sortierte Liste

1. Wiederhole $(n - 1)$ - mal:
2. Für jedes i von 1 bis $n - 1$:
3. Vergleiche das Objekt X_i mit seinem Nachbarn X_{i+1} , falls X_i größer als X_{i+1} ist:
4. Tausche X_i und X_{i+1}

Nach diesem Durchlauf (oder *Iteration*) sind die Platten schon etwas besser sortiert (NIR ist die letzte), aber noch nicht in der richtigen Reihenfolge. Nach zwei weiteren Iterationen befindet sich die Liste in der richtigen Reihenfolge – wie in Abb. 2:

- a) DTH kommt vor MAN und die Platten bleiben in dieser Reihenfolge: DTH – MAN – DJA – NIR