

Simulation und Modellierung - Verständniszusammenfassung

16. Februar 2016

1 Spieltheorie

Hier widmen wir uns Spielen zwischen zwei Spielern. Um über Spiele mathematisch nachdenken zu können, führen wir Notationen ein. Dann teilen wir die Spiele in zwei Kategorien ein: Eine Kategorie sind Spielen bei denen ein Spieler unabhängig vom anderen seine Entscheidung trifft (Glücksspiel etc.). Die andere Kategorie sind Spiele, bei denen sich ein beide Spieler Gedanken darüber machen, was der andere tut (z.B. Gefangenendilemma). Dabei behandeln wir von beiden Spielarten die Strategien der Spieler. Am Ende gehen wir kurz auf Nash Gleichgewichte ein.

1.1 Spiele in strategischer Normalform

Wichtige Begriffe:

- Spieler
- Menge der Strategien von einem Spieler X
- Menge aller Strategiepaare: jeder Strategie vom Spieler A werden alle Strategien vom Spieler B zugeordnet und andersrum
- Auszahlungsfunktion: Diese beschreibt den Nutzen eines Strategiepaares für einen Spieler
- Nutzenmatrix: Die Auszahlungsfkten beider Spieler für alle Strategien werden in einer Matrix zusammengefasst. Dies ermöglicht eine genaue Übersicht über die Gewinne/ Verluste

1.2 Spiele ohne Annahmen über den Gegner

Hier trifft ein Spieler seine Entscheidung, ohne sich Gedanken über den anderen Spieler zu machen. Ein Beispiel ist das Glücksspiel bzw. ob ein Spieler einen Regenschirm mitnimmt oder nicht.

1.2.1 Spiel bei Gewissheit

Falls ein Spieler alle Strategien vom anderen Spieler kennt, kann er einfach seinen Nutzen maximieren.

1.2.2 Spiel bei Risiko

Hier hat der Spieler keine Information, wie der andere Spieler spielen wird. Er muss sich entscheiden, wie er sich verhält. Dabei gibt es den risikobereiten und den vorsichtigen Spieler. Der Risikobereite Spieler geht davon aus, dass der andere Spieler seinen Gewinn maximieren möchte und maximiert auf dieser Annahme seinen Gewinn. Der vorsichtige Spieler jedoch versucht den garantierten Gewinn zu maximieren.

$$U = \begin{pmatrix} 0 & 30 \\ 10 & 10 \end{pmatrix}$$

bei diesem Beispiel würde der risikobereite Spieler die erste Zeile wählen und der vorsichtige die zweite Zeile.

1.3 Reaktionsabbildungen

Nun widmen wir uns dem Spiel, bei dem beide Spieler ihren Gewinn maximieren wollen. Also muss jeder Spieler überlegen, was er tun würde, wenn der Gegner eine bestimmte Strategie wählt. Dazu ist die Reaktionsabbildung gut. Sie bildet jede Strategie vom Gegner auf die Strategie ab, die für ihn die optimale Antwort auf die Strategie des Gegners ist.

Die Gesamtreaktionsabbildung bildet jedes Strategiepaar auf die optimale Antwort des Gegners ab. Also wie reagiert Spieler A auf die Strategie b von Spieler B und wie reagiert Spieler B auf die Strategie a von Spieler A. Diese Antwort kann in der Nutzenmatrix durch unterstreichen markiert werden.

1.4 Dominante Strategien

Auch hier überlegen wieder beide Spieler gleichzeitig und beziehen die Überlegung des anderen mit ein.

Falls ein Spieler eine Strategie besitzt, die auf alle Strategien vom anderen Spieler die beste Antwort ist, kann er sie ohne Bedenken spielen. Dies ist die dominante Strategie. Also wird eine Strategie gewählt, die den Spieler unabhängig vom Gegner macht.

Eine Strategie nennen wir redundant, falls sie durch eine andere Strategie dominiert wird. Bei Schere- Stein-Papier-Brunnen ist der Stein redundant und kann von beiden Spieler weggelassen werden.

Leider ist nicht immer eine dominante Strategie verfügbar.

1.5 Nash-Gleichgewicht

Ein Nash Gleichgewicht liegt vor, wenn Strategien a und b (von Spieler A bzw. B) existieren, sodass die optimale Antwort von B auf a die Strategie b ist und die optimale Antwort von A auf Strategie b die Strategie a ist. Dabei ist ein Strategiepaar aus dominanten Strategien immer ein Nash Gleichgewicht.

Es wäre schön, wenn es immer nur einen Nash Gleichgewichtspunkt geben würde. Dem ist aber leider nicht so.

1.6 Gemischte Strategien

In diesem Kapitel betrachten wir nur Nullsummenspiele, also Spiele bei denen die Strategien des einen Spieler genau die entgegengesetzten Strategien des anderen Spielers sind.

Bei gemischten Strategien wählen beide Spieler ihre Strategien mit einer bestimmten Wkeit. Diese Wkeit nennt man gemischte Strategien.

Die Auszahlungsfkt lässt sich dabei berechnen als Multiplikation von dem Wkeitsvektor der Strategie des einen Spielers mit der Nutzenmatrix mit dem Wkeitsvektor der Strategie des anderen Spielers.

Frage: Warum lässt sich das so berechnen?

Dabei versuchen beide Spieler ihren Wkeitsvektor so zu wählen, dass der Erwartungswert für ihn maximal wird.

Es lässt sich beobachten, dass falls die Nutzenmatrix nicht invertierbar ist ein Wkeitsvektor existiert, sodass dieser Wkeitsvektor mit der Nutzenmatrix multipliziert 0 ergibt, dass dann der Spieler diesen Wkeitsvektor völlig unabhängig vom anderen Spieler wählen.

Mathematische Herleitung ist klar, aber leider nicht die Interpretation.

Ein Beispiel dafür ist die das Spiel Schere Stein Papier. Beide Spieler sollten Gleich verteilt spielen.

1.7 Ausblick

Mathematisch lässt sich diese Modellierung sehr gut fassen und die meisten Spiele lassen sich mathematisch sehr viel einfacher durchdenken, als in Prosa Form. Das Problem (wie immer in der Modellierung) ist die Wahl der Auswertungsfunktion, also insbesondere die Werte in der Matrix und der Abschätzung was dem Spieler wichtiger ist. Z.b. könnte ihm der reine Gewinn wichtig sein oder das Ärgernis über entgangenen Gewinn.

2 Gruppenentscheidungen

Wie der Titel andeutet haben wir in dieser Kategorie eine Menge von Wählern die Kandidaten in eine bestimmte Reihenfolge bringen sollen. Die daraus resultierende Gesamtreihenfolge soll die Wahl der Wähler möglichst gut repräsentie-

ren. Was dieses möglichst gut heißt, werden wir genauer definieren und merken, dass es kein perfektes Verfahren gibt, dass alle Stimmen gut repräsentiert.

2.1 Individualpräferenzen und Gruppenentscheidungen

Um Gruppenentscheidungen treffen zu können, benötigen wir eine Menge von Wählern und eine Menge von Kandidaten. Dabei soll jeder Wähler jedem Kandidaten einen Rang zuordnen, also eine natürliche Zahl. Je kleiner diese ist, desto lieber hat der Wähler den Kandidat. Dies macht die Rangabbildung. Sie ordnet jedem Kandidat eine natürliche Zahl zu. Sie muss surjektiv sein.

Frage: Warum surjektiv? Antwort: Falls ich genauso viele Ränge, wie Kandidaten habe, kann ich keinen zwei Kandidaten den gleichen Rang zuordnen. Falls ich weniger Ränge habe, dann bin ich gezwungen mehrere Kandidaten gleich hoch zu wählen. Ergibt diese Forderung wirklich Sinn? Wobei wir im Folgenden auch mal damit arbeiten, dass wir zwei Kandidaten den gleichen Rang zuordnen...

Diese Abbildung definiert eine transitive und asymmetrische Präferenzrelation. Damit zwei Kandidaten in Relation zueinander stehen, muss der Wähler einen Kandidaten strikt besser finden, als den anderen. Sonst haben die Kandidaten keine Relation.

Damit diese nicht mehr passiert, gibt es die Präferenzrelation Stern. Bei dieser stehen auch zwei Kandidaten in Relation, wenn sie den gleichen Rang haben.

Die Präferenzrelation und die Präferenzrelation Stern sind invers komplementär zueinander. d.h. falls entweder die Präferenzrelation oder die Präferenzrelation Stern gegeben ist, kann daraus die jeweils andere bestimmt werden.

Nun wollen wir die Präferenz der Gesamtheit bestimmen. Diese wird durch eine Funktion berechnet, die sogenannte Wohlfahrtsfkt oder kollektive Auswahlfkt, die n Präferenzen der Wähler als Eingabe bekommt und daraus die gesamte Präferenz berechnet.

Wir wollen Bedingungen an die kollektive Auswahlfkt stellen:

Die Abbildung muss total sein. D.h. jede beliebige Präferenz als Eingabe muss eine Ausgabe haben.

Außerdem muss das Ergebnis der kollektiven Auswahlfkt ebenfalls eine Relation sein, die jedem Kandidaten einen eindeutigen Rang zuordnet.

Frage: Warum verlangen wir einen Eindeutigen Rang? Man könnte auch zwei gleich gute/ schlechte Kandidaten haben. Vll gibt es dann keinen Sieger der Wahl?

2.2 Beispiele für Entscheidungsverfahren

In diesem Kapitel werden ein paar Entscheidungsverfahren vorgestellt. Wie wollen probieren möglichst gerechte Verfahren zu konstruieren, wobei gerecht im

Philosophischen bzw. juristischen Sinn gemeint ist.

Externer Diktator

Der Titel ist klar: Ein Mensch, der kein Wähler ist, bestimmt den Ausgang der Wahl. Mathematisch hat die kollektive Auswahl fkt egal bei welcher Eingabe, die Ausgabe, die der Diktator möchte.

Interner Diktator

Hier bestimmt ein Wähler, welches Ergebnis bei der Wahl raus kommt. Also Der Ausgang ist immer durch den i-ten Wähler genau festgelegt, egal, was die anderen wählen.

Rangaddition

Die Präferenz der Wähler gibt eine Rangabbildung vor.

Frage: Ist diese Rangabbildung eindeutig? Ja, weil wir Surjektivität gefordert haben?

Bei der Rangaddition werden alle Ränge für jeden einzelnen Kandidaten addiert. Nun wird wieder geschaut, in welcher Reihenfolge die Kandidaten sind.

Das Problem hier ist, dass die aus der Rangaddition resultierende Rangabbildung nicht surjektiv ist. Bsp: Zwei Kandidaten, zwei Wähler. Beide finden Kandidat A besser, also hat Kandidat A 2 und Kandidat B 4 Pkt. Die einfache Lösung für das Problem ist, dass man die bijektiv abbildet, also die 2 auf die 1 und die 4 auf die 2.

Cordoc-Verfahren

In diesem Verfahren vergleicht man immer zwei Kandidaten miteinander. Man nehme zwei Kandidaten und alle Präferenzrelationen der n Wähler und zählt die Vergleiche, die Kandidat A und die Kandidat B gewinnt. Gewinnt A mehr Vergleiche, so ist Kandidat A vor Kandidat B. Dies tut man nun für alle Kandidaten.

Das Problem bei dieser Wahl ist, dass Zyklen entstehen können. Es kann passieren, dass Kandidat A von B vor C vor A ist. Das wollen wir natürlich nicht.

Einstimmigkeit

Ein Kandidat x wird nur vor y gesetzt, falls alle Wähler x besser finden als y.

Dieses Verfahren ist kaum anwendbar, da so was eig nie passiert. Außerdem muss das Einstimmigkeitsprinzip keine Relation liefern, da manche Kandidaten keinem Rang zugeordnet werden.

Borda-Wahl

Jeder Wähler gibt den k Kandidaten Punkte. Dabei bekommt der erste Wähler k Punkte, der zweite $k-1$ usw. Der Kandidat mit den meisten Pkt gewinnt die Wahl.

Frage: Was ist der unterschied zur Rangaddition? Antwort: Bei der Rangaddition ist am Anfang zugelassen, dass zwei Kandidaten die gleiche Anzahl an Pkt bekommt.

Hier gibt es aber auch ein Problem: Wir haben eine Wahl und Kandidat A ist auf letzter Stelle gekommen. Nehmen wir die gleichen Wahlergebnisse nochmal, nur dass A nicht antritt, kann es sein, dass ein anderer erster wird.

Insgesamt zeigt sich, dass die Auswahl des Wahlverfahrens eine große Rolle dabei spielt, was für ein Ergebnis am ende raus kommt.

2.3 Bedingungen an die Auswahlfkt und Satz von Arrow

Wir wollen nun genauer definieren, was ein gerechtes Verfahren ist.

Um ein gerechtes Verfahren zu erhalten, muss die Pareto Bedingung erfüllt sein. Das heißt, dass falls sich alle Wähler einig sind, muss jedes mögliche Wahlergebnis erreichbar sein.

Außerdem sollte die Unabhängigkeit von irrelevanten Ereignissen erfüllt sein. Die Reihenfolge von zwei Kandidaten sollte nicht dadurch geändert werden können, dass Wähler ihre Präferenz bzgl. eines dritten Kandidaten verändern.

Es zeigt sich leider im Satz von Arrow, dass kein Verfahren, außer der Interne Diktator alle Bedingungen an ein Verfahren erfüllt. Anders formuliert: Es gibt keine kollektive Auswahlfkt, die alle fünf demokratischen Grundregeln erfüllt.

Die Demokratischen Grundregeln sind:

- die Auswahlfkt muss total sein.
- Das Ergebnis muss eine Relation sein, die einer Rangabbildung genügt.
- Die Pareto bedingung muss erfüllt sein
- die Unabhängigkeit von irrelevanten Alternativen muss erfüllt sein
- es gibt keinen Diktator

2.4 Nicht-Manipulierbarkeit von Wahlen

Die soziale Entscheidungsfkt bildet die Menge der Präferenzen von den Wählern auf die Menge der Kandidaten ab. Also wird jede Präferenz aller Wähler genau ein Kandidat zugeordnet. Dieser Kandidat ist der Sieger der Wahl.

Die soziale Entscheidungsfkt heißt strategisch manipulierbar, falls ein Wähler der den Kandidaten B vor den Kandidaten A präferiert, statt seiner Präferenz eine andere angeben kann und damit den Kandidaten B erzwingen kann.

Die soziale Entscheidungsfkt heißt Anreizkompatibel, falls sie nicht manipulierbar ist.

Eine andere Definition von anreizkompatibel ist: falls nur ein Wähler seine Meinung ändert und bei beiden Meinungen unterschiedliche Kandidaten (beim ersten mal a, beim zweiten b) herauskommen, dass man dann schon schlussfolgern kann, dass bei der ersten Wahl der Wähler a vor b gesetzt hat und bei der zweiten genau andersherum.

Nun stellt sich aber folgendes Problem dar: Sobald eine Entscheidungsfunktion anreizkompatibel und surjektiv ist, so ist sie schon diktatorisch. Dies ist der Satz von Gibbard-Satterthwait.

3 Informationssuche im Netz

Ziel dieses Kapitels ist es, eine Idee davon zu bekommen, wie Suchmaschinen die Ergebnisse von einer Suchanfrage ermittelt. Dabei wollen wir eine möglichst akzeptable Lösung der Problemstellung mit den gewünschten Eigenschaften bekommen.

Die Idee ist es, dass wir einen Index mit allen relevanten Webseiten haben. Zu jeder Webseite listen wir alle möglichen Suchbegriffe auf. Dieses Verfahren ist die Suche im Index.

Die Suche im invertierten Index ist nun, dass zu jedem möglichen Suchbegriff alle relevanten Webseiten aufgelistet werden. Diese Webseiten müssen nun anhand einer Ratingfunktion gewichtet werden. Dabei sollte die Relevanz und Qualität einer Webseite eine Rolle spielen, ebenso wie die Sichtbarkeit des Suchbegriffes.

Das Besondere in diesem Kapitel ist es, dass die Mathematik selber die Lösung für das Problem ist. Der Begriff des Page Rank Vektors ist von zentraler Bedeutung. Er beschreibt die Wichtigkeit der einzelnen Seiten. Zunächst erstellen wir ein Modell, bei dem ein websurfer von einer Seite zur nächsten springt, indem er einen Link auf einer Seite folgt. Hier ist das Problem, dass er nichts tun kann, wenn kein Link vorhanden ist. Also verhält sich im nächsten Modell der Websurfer so, dass er, falls kein Link da ist, gleichverteilt auf eine beliebige Seite springt. Hier ist das Problem, dass es Zyklen geben kann, aus denen der Websurfer nicht mehr herauskommt. Also kann er in jedem Schritt mit einer bestimmten Wkeit, auch wenn Links vorhanden sind, zu einer beliebigen Seite springen. Dieses letzte Modell ist das beste von allen, die wir kennenlernen. Diese Modelle kann man als Eigenwert problem umschreiben. Die Existenz und Eindeutigkeit einer Lösung bestimmen wir mittels Markov Ketten und der Perron Frobenius Theorie. Schließlich wollen wir auch noch die Lösung des Eigenwert Problems bestimmen. Dies können wir durch Vektoriteration bzw. lösen eines Gleichungssystems.

3.1 Page-Rank und Web-Graph

Wann ist ein Artikel nun wichtig? Die Grundlegende Idee ist, dass eine Webseite wichtig ist, wenn viele andere Webseiten ihn Zitieren und deswegen Links zu dem Artikel haben.

Also können wir einen Graphen zeichnen, deren Knoten die Webseiten und die Kanten die Links/ Verweise sind.

Dazu definieren wir den Page-Rank-Vektor, der die Wichtigkeit der einzelnen Webseiten in einem Vektor darstellt.

Als erstes Modell nehmen wir ein lineares: Dabei ist der Page-Rank von einer Seite

$$r(S_i) = \sum_{S_j \in B_{S_i}} \frac{r(S_j)}{|S_j|} \quad (1)$$

wobei B_{S_i} die Menge aller Seiten ist, die einen Link auf S_i haben. $|S_j|$ die Anzahl der Seiten ist, auf die die Seite S_j zeigt und $r(S_j)$ der Pagerank von der Webseite S_j ist.

Man kann iterativ den Pagerankvektor bestimmen.

To do: Beschreiben, wie es geht, falls zeit

Betrachten wir nun die Hyperlinkmatrix von einem Web Graphen. Der ij -te Eintrag ist das Inverse der Anzahl der Links, die vom Graphen i ausgehen, falls die Webseite S_i einen Link auf die Seite S_j hat und 0 sonst.

Damit kann (1) als Eigenwertproblem ausgedrückt werden: Der Pagerankvektor ist das gleiche wie die transponierte Hyperlinkmatrix mal dem Pagerankvektor. Also wird der Pagerankvektor, der der Eigenvektor zum Eigenwert 1 ist, gesucht.

Dabei erhält man leider nicht zu jedem Webgraphen eine eindeutige Lösung. Dieses Problem entsteht, wenn eine Webseite existiert, die auf keine anderen Webseiten zeigt.

Der nächste Abschnitt zeigt, wie man mit diesem Problem umgehen kann.

3.2 Zufallssurfer und Markovketten

Ein Websurfer geht zufällig auf eine Seite und sucht sich dort zufällig einen Link aus, den er folgt. Der Page Rank ist nun die Wkeit, mit der sich ein Nutzer auf einer Seite befindet.

Über den Zufallssurfer treffen wir nun die Annahme, dass er gedächtnislos ist und falls Links vorhanden sind, er einen Link zufällig (gleichverteilt) folgt oder falls kein Link vorhanden ist, auf eine beliebige Seite springt.

3.2.1 Markovketten

Um die bessere Modellierung für das Google Problem zu bekommen brauchen wir ein wenig Markovketten Theorie.

Der Zustandsraum E ist die Menge aller möglichen Zustände. Bei uns sind das alle identifizierten Webseiten.

Ein stochastischer Prozess ist eine Abbildung $X : T \rightarrow E$ die jedem Zeitpunkt einen Zustand zuordnet, also $X(t) = X_t$. Bei uns heißt das, dass ein Nutzer zu einer bestimmten Zeit auf einer bestimmten Webseite ist. Wir brauchen nur diskrete Zeitzustände.

$\mathbb{P}[X_t = x]$ ist die Wkeit, dass zu einem Zeitpunkt t der Zustand x vorliegt, also hier die Wkeit, dass ein Nutzer zum Zeitpunkt t auf der Webseite x ist.

Häufiger werden wir die bedingte Wkeit $\mathbb{P}[A|B]$ gebrauchen, die die Wkeit angibt, dass A eintritt unter der Voraussetzung, dass B eingetreten ist. Dies wird berechnet durch. $\mathbb{P}[A|B] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

Wir gehen davon aus, dass wir zum k -ten Zeitpunkt einen bestimmten Zustand erreicht haben. Nun wollen wir die Wkeit für ein bestimmtes Ereignis zum nächsten Zeitpunkt bestimmen. Falls diese Wkeit nicht davon abhängig ist, welche Ereignisse vor dem k -ten Zeitpunkt eingetreten sind, dann nennen wir den stochastischen Prozess einen Markowprozess. Wenn wir die Webseiten als Ereignisse und die diskreten Zeitpunkte betrachten ist das wechseln von einer Webseite zu einer anderen ein Markowprozess, da er nicht davon abhängt, welche Seiten der Nutzer vorher besucht hat. Die Links auf einer Seite verändern sich dadurch nicht.

Frage: wie kann ich homogenen Markowprozess erklären?

Der i -te Eintrag des Zustandswahrscheinlichkeitsvektor $\pi^{(k)} \in \mathbb{R}^n$, wobei n die Anzahl der möglichen Zustände sind, gibt die Wkeit an, dass zum Zeitpunkt k der Zustand i eingetreten ist. Dies ist der Pagerank Vektor zum Zeitpunkt k .

3.2.2 Anwendung auf den Zufallssurfer

Wenden wir nun diese Theorie an, kann man die Übergangsmatrix U der Markowkette aufstellen. Der ij -te Eintrag ist, wie vorher das Inverse der Anzahl der Links, die vom Graphen i ausgehen, falls die Webseite S_i einen Link auf die Seite S_j hat oder das Inverse der Anzahl aller Webseiten, falls die Webseite i überhaupt keinen Link hat und 0 sonst.

Die Matrix U ist stochastisch, da die Summe der Einträge einer Zeile 1 ist und kein Eintrag negativ ist.

Aus der Markoweigenschaft folgt nun, dass $\pi^{(k)T} = \pi^{(k-1)T}U = \dots \pi^{(0)T}U^k$ gilt.

Frage: Warum gilt das?

Mit diesem Wissen wollen wir nun den Pagerank Vektor π bestimmen, für den gelten soll, dass $\pi^T = \pi^T U$. Außerdem wollen wir, dass ein paar Eigenschaften für π gelten.

Wir wollen zunächst, dass der Pagerank die mittlere Aufenthalts Wkeit eines Nutzers auf einer Internetseite angibt. Außerdem soll, falls wir den Markowpro-

zess schon mit dem berechneten Pagerank starten, wieder der Pagerank zu jedem Zeitschritt rauskommen. Zum Schluss soll noch gelten, dass, wenn wir mit einer beliebigen Anfangsverteilung starten, der Markowprozess gegen den Pagerankvektor strebt.

Wir suchen also den Eigenvektor zum Eigenwert 1 der Gleichung $\pi = U^T \pi$.

Nun wollen wir mit der Perron-Frobenius Theorie herausfinden, ob wir einen Pagerankvektor finden, der all diese Eigenschaften genügt. Auch die Eindeutigkeit ist von großer Bedeutung.

3.3 Perron-Frobenius Theorie

In diesem Kapitel gehe ich nicht auf die mathematischen Grundlagen, sondern nur auf die wichtigsten Ergebnisse ein.

Wir wollen das asymptotische Verhalten von $x^{k+1} = Ux^k$ untersuchen.

Zunächst bekommen wir leider nur eine Aussage durch den Satz von Perron über eine positive Übergangsmatrix. Bei dieser existiert die stationäre Verteilung und ist eindeutig. Außerdem können wir sagen, dass der Grenzwert der Folge $\pi^k = \pi$ ist.

Leider lässt sich das Theorem noch nicht auf die Übergangsmatrix des Zufallssurfers anwenden, da U nur nicht negativ ist. Falls wir nun annehmen, dass U auch irreduzibel ist, dann finden wir einen einfachen Perron Eigenwert $\lambda = 1$ und einen zugehörigen Linkseigenvektor $\pi^T > 0$. Nun haben wir aber nicht, dass dies der einzige positive Eigenwert ist.

Frage: Wo brauchen wir primitive Matrizen?

Wir wollen ausschließen, dass der Web Graph zyklert und einige Zustände nicht mehr erreichen kann. Deshalb verstärken wir unsere Annahmen an den Zufallssurfer: Dieser findet mit einer Wkkeit a eine neue Seite, indem er zufällig einem Link auf der momentanen Seite folgt. Falls kein Link vorhanden ist, auf eine beliebige andere Webseite springt und mit einer Wkkeit von $1-a$ zufällig auf eine beliebige Seite springt. Dadurch erhalten wir die Google Matrix. Diese ist irreduzibel, aperiodisch und stochastisch. Also hat sie eine eindeutig stationäre Grenzverteilung π . Also hat das Problem $G^T \pi = \pi$ mit $\pi^T e = 1$ eine eindeutige Lösung.

3.4 Lösungsstrategien und Sensitivitätsanalyse

Betrachten wir dieses Problem genauer: G ist gegeben durch $G = \alpha U + (1 - \alpha) \frac{1}{n} ee^T$ wobei $U = H + \frac{1}{n} ae^T$. Dabei ist

- H Hyperlinmatrix - dünn besetzt
- U Übergangsmatrix, dünn besetzt, stochastisch
- G Google Matrix - voll besetzt, primitiv, stochastisch
- $E = \frac{1}{n} ee^T$ Teleportationsmatrix

- n Anzahl der Webseiten - sehr groß
- α Skalierungsparameter
- π^T stationäre Verteilung als Zeilenvektor
- a binär hängender Knotenvektor

Betrachten wir nun das Eigenwertproblem und setzen dort die Definition für G ein. Wir erhalten nach umgerechnen: $(I - \alpha U^T)\pi = \frac{1-\alpha}{n}e$. Da n zu groß ist, ist das direkte Verfahren ungeeignet

Frage: Später haben wir immer noch n . Warum ist es dann besser?

Deshalb wenden wir ein iteratives Verfahren an: $\pi^{(k+1)} = G^T \pi^{(k)}$. Da G aber voll besetzt ist, ist der Rechenaufwand zu groß. Deshalb ersetzen wir wieder G durch seine Definition und haben dann das Verfahren $\pi^{(k+1)} = \alpha H^T \pi^{(k)} + \frac{e(\alpha a^T \pi^{(k)}) + 1 - \alpha}{n}$

Bis jetzt haben wir nur als Anfangsverteilung die Gleichverteilung betrachtet (e^T in der Google Matrix). Diese wollen wir nun durch eine beliebige Anfangsverteilung ν^T ersetzen. Daraus ergibt sich: $(I - \alpha U)\pi^T = (1 - \alpha)\nu^T$

Potenzmethode

Sei eine Matrix A gegeben, wobei ein Eigenwert betragsmäßig echt größer ist, als die anderen Eigenwerte und ein Anfangswert x_0 . Dann berechnet die Potenzmethode den Betragsmäßig größten Eigenwert und den dazugehörigen Eigenvektor. Genau dies wollen wir für die Google Matrix erreichen. Dann berechnet sich dies durch: $\hat{x}_k = Ax_{k-1}$ wobei $x_k = \frac{\hat{x}_k}{\max(|\hat{x}_k|)}$.

Es gibt auch andere, „bessere“ Methoden die auch den gewünschten EW und zugehörigen EV berechnen, aber die Potenzmethode ist einfacher zu implementieren und bei anderen Methoden muss ein Vektor zwischengespeichert werden, dies ist aber nicht möglich, da n zu groß ist.

3.4.1 Konvergenzgeschwindigkeit und Wahl des Parameters α

Bekannt ist, dass die Konvergenzgeschwindigkeit der Potenzmethode folgender Abschätzung genügt: $\|\pi^{(k)} - \pi\|_1 \leq c \frac{|\lambda_2|^k}{|\lambda_1|^k}$, wobei die λ_i die Eigenwerte sind und der Betragsmäßig größte EW λ_1 ist, dann λ_2 usw.

Wir können beweisen, dass das Spektrum der Google Matrix das Gleiche ist, wie das der Übergangsmatrix, bis auf dem Vorfaktor α .

Normalerweise ist λ_2 der Übergangsmatrix ungefähr 1. Dies führt dazu $\lambda_2 = \alpha$ bei der Google Matrix gilt. Also: Je kleiner $\alpha > 0$ ist, desto schneller konvergiert die Potenzmethode, aber desto mehr widerspricht das der Modellierungs-idee. Es zeigt sich, dass $\alpha = 0,85$ eine gute Wahl ist.

Die bisherigen Überlegungen in Bezug auf den Fehler bezogen sich nicht auf die Reihenfolge in π . Dies ist fragwürdig, da genau dies wichtig ist. Deshalb könnte man noch andere Fehlermaße betrachten.

3.5 Sensitivitätsanalyse

In diesem Kapitel gehen wir der Frage nach, wie empfindlich der Page Rank Vektor auf Änderungen von dem Parameter α , H und der Anfangsverteilung ν^T reagiert.

Sensitivität bzgl α

Um die Sensitivität, also Veränderungen bzgl α festzustellen, leiten wir den Pagerank Vektor nach α ab. Es ergibt sich, dass $\left| \frac{d\pi_j(\alpha)}{d\alpha} \right| \leq \frac{1}{1-\alpha}$ und deswegen

$$\left\| \frac{d\pi(\alpha)}{d\alpha} \right\| \leq \frac{2}{1-\alpha}$$

Frage: Müsste statt der 2 da kein n stehen?

Falls α nicht zu nahe bei 1 liegt, reagiert der Page Rank Vektor nicht all zu sensitiv auf Änderungen. Für $\alpha \rightarrow 1$ wird die Abschätzung unbrauchbar. Es zeigt sich sogar, dass der Pagerank Vektor immer empfindlicher auf Änderungen reagiert, je näher α bei 1 liegt.

Betrachten wir nun ein Update der Google Matrix und den daraus resultierenden neuen Page Rank Vektor $\tilde{\pi}^T$. Dann gilt $\|\pi - \tilde{\pi}\|_1 \leq \frac{2\alpha}{1-\alpha} \sum_{i \in U} \pi_i$, wobei U die Menge aller Seiten ist, die aktualisiert wurden. Diese Abschätzung ist hilfreich, falls α klein ist und die Summe des Page Ranks ebenfalls.

Sensitivität bzgl H

Diese Sensitivitätsanalyse ist besonders wichtig, da sich das Internet ständig verändert. Man kann zeigen, dass $\frac{d\pi^T(h_{ij})}{dh_{ij}} = \alpha\pi_i(e_j^T - \nu^T)(I - \alpha U)^{-1}$ gilt. Der Term π_i zeigt, dass Veränderungen auf wichtigen Seiten mehr Einfluss haben als Veränderungen auf unwichtigen.

Sensitivität bzgl ν^T

Betrachten wir zunächst ν genauer. Jeder Nutzer von Google hat unterschiedliche Präferenzen, also personalisieren wir ν , indem wir für ν unterschiedliche Kategorien einführen wie Kunst, Wirtschaft, Fernsehen etc. Der Pagerank Vektor hängt nun mit unterschiedlichen Gewichten von diesem Personalisierten Anfangswerten ab. Also $\pi^T = \beta_1 \pi^T(\nu_1^T) + \dots + \beta_{16} \pi^T(\nu_{16}^T)$ mit $\sum_{i=1}^{16} \beta_i = 1$

Betrachten wir nun die Sensitivität bzgl ν . Man kann zeigen, dass $\frac{d\pi^T(\nu^T)}{d\nu^T} = (1 - \alpha \sum_{i \in D} \pi_i) (I - \alpha U)^{-1}$ wobei D die Menge der hängenden Knoten ist.

Frage: Bewertung der Formel?

3.6 Page Rank Vektor Berechnung - andere Wege

Es gibt auch noch andere Wege, den Page Rank Vektor zu berechnen. Z.B. kann man das Gleichungssystem $x^T(I - \alpha U) = \nu^T$ nach x auflösen um dann den Pagerankvektor $\pi^T = \frac{x^T}{x^T e}$ zu erhalten.

Für kleine Systeme ist das Lösen auf dieser Art besser und vor allem einfacher. Desweiteren ist die Potenzmethode hinsichtlich der Konvergenz empfindlicher für $\alpha \rightarrow 1$. Das LGS ist es ist. Aber auch hier ist der Page Rank Vektor empfindlich für $\alpha \rightarrow 1$.

Eine andere Art ist es, die Google Matrix irreduzibel zu gestalten, indem ein Dummy Knoten eingebaut wird, sodass der Zufallssurfer von jeder Webseite aus mit einer Wkeit $(1 - \alpha)$ zu diesem Dummy Knoten kommt und sich dann mit der Wkeit der Anfangsverteilung auf die anderen Webseiten bewegt. Diese Matrix sieht dann so aus:

$$\hat{U} = \begin{pmatrix} \alpha U & (1 - \alpha)e \\ \nu^T & 0 \end{pmatrix}$$

Man kann weiterhin zeigen, dass das Spektrum vom Spektrum der Übergangsmatrix U abhängt: Das Spektrum von U sei gegeben durch $\{1, \lambda_1, \dots, \lambda_n\}$. Dann ist das Spektrum von \hat{U} gegeben durch $\{1, \alpha\lambda_1, \dots, \alpha\lambda_n, \alpha - 1\}$.

4 Verkehrssimulation

Es gibt makroskopische und mikroskopische Verkehrssimulation. Die Makroskopische beschäftigt sich mit dem Verkehr im ganzen und lässt einzelne Fahrzeuge außer acht und die mikroskopische betrachtet die einzelnen Fahrzeuge.

4.1 Makroskopische Verkehrssimulation

Beginnen wir mit der makroskopischen Verkehrssimulation.

Zur Betrachtung des Verkehrs brauchen wir einige Größen: Die Verkehrsdichte beschreibt, wie dicht die Fahrzeuge auf der Straße stehen, also wie viele Fahrzeuge pro Km auf der Straße sind. Der Verkehrsfluss beschreibt, wie viele Fahrzeuge pro Zeiteinheit einen Kontrollpunkt passieren.

Frage: Wofür wollen wir ein Modell? Was wollen wir damit aussagen über den Verkehr?

Zunächst treffen wir Annahmen über den Verkehr. Einmal ist die Fahrzeugdichte begrenzt durch eine minimale Kapazität 0 und eine maximale Kapazität, also Stoßstange an Stoßstange. Außerdem dürfen die Fahrzeuge nicht rückwärts fahren, wir befinden uns auf einer Autobahn, und die Autos haben eine maximale Geschwindigkeit

Frage: wo wird die wichtig?

Desweiteren nehmen wir an, dass der Verkehrsfluss stetig differenzierbar ist.

Verkehrsfluss, Dichte und Geschwindigkeit stehen im Zusammenhang. Dimensionsanalyse ergibt, dass Fluss = Dichte mal Geschwindigkeit gelten muss. Dies ergibt Sinn:

Betrachten wir ein Auto auf einer Straße, bei der die Dichte ρ_0 konstant ist. Angenommen das Auto fährt T Stunden. Dann kann es mit konstanter Geschwindigkeit v_0 $T \cdot v_0$ Kilometer zurücklegen. Alle anderen Autos auf dieser Straße können dies ebenfalls. Also kommen an einem Kontrollpunkt in T Stunden $T \cdot v_0 \cdot \rho_0$ Autos vorbei. Also ist der Verkehrsfluss $\rho_0 v_0$ Autos pro Stunde. Da konstante Geschwindigkeit und konstante Verkehrsdichte nicht realistisch ist, betrachtet man nur ein sehr kleines Zeitintervall und kommt wieder auf die Formel $q = \rho \cdot v$.

Betrachten wir eine unendlich lange Straße. Da die Anzahl der Autos auf der Straße sich nur durch das Verlassen/ Ankommen auf der Straße ändert, verändert sich die Anzahl der Autos auf dieser unendlich langen Straße nicht, bleibt also konstant. Dies ist der Integrale Erhaltungssatz.

Die Herleitung findet sich in Formeln im Script.

Der differentiale Erhaltungssatz besagt, dass die Veränderung der Dichte in der Zeit das Gleiche ist, wie die negative Veränderung des Flusses im Ort. Besagt also: Wenn sich die Dichte der Autos verändert, also mehr (weniger) Fahrzeuge im Laufe der Zeit am gleichen Ort befinden, so wird der Fluss entsprechend niedriger (höher), also passieren weniger (mehr) Autos den Kontrollpunkt.

Da sich der Fluss als Geschwindigkeit mal Dichte beschreiben lässt, wollen wir eine Formel für die Geschwindigkeit herleiten.

Machen wir zunächst ein paar Annahmen über diese.

Die Geschwindigkeit hängt von der Dichte ab. Ist die Straße leer, so kann mit voller Geschwindigkeit gefahren werden. Ist die Straße befahren, so sollte langsamer gefahren werden, als wenn sie leer ist, also je voller die Straße desto langsamer fährt man. Ist die Straße komplett voll, so ist die Geschwindigkeit 0.

Dabei kommt der Verkehr zum Erliegen, bevor die Fahrzeuge sich berühren, bzw wenn gar keine Autos auf der Straße sind.

Nun kann man mittels der eigenen Beschleunigung, die von der Geschwindigkeit des Vorfahrers abhängt, die Geschwindigkeit ermitteln. Die Herleitung mit Formeln findet sich im Script. Wir erhalten:

$$v(x, t) = \begin{cases} v_{max} & \rho \leq \rho_{crit} \\ \lambda \left(\frac{1}{\rho(x, t)} - \frac{1}{\rho_{crit}} \right) & \rho > \rho_{crit} \end{cases}$$

Diese Formel bedeutet nun: Falls weniger Autos, als eine kritische Dichte auf der Straße sind, kann die maximale Geschwindigkeit gefahren werden. Falls dem nicht so ist, gilt: Je mehr Autos auf der Straße sind, also je näher die momentane Dichte an die maximale Dichte herankommt, desto langsamer fahren die Autos.

Dies muss aber noch mit der Sensitivität multipliziert werden.

Frage: Warum?

Bis jetzt haben wir angenommen, dass die Sensitivität konstant ist. Dies ist jedoch nicht der Fall. Sie ist davon abhängig, wie weit die Autos voneinander entfernt sind. Je weiter die Autos voneinander entfernt sind, desto kleiner ist die Sensitivität (Störungen?) und je dichter die Autos beieinander sind, desto größer ist sie. Damit ergibt sich eine andere Formel für die Geschwindigkeit.

$$v(x, t) = \begin{cases} v_{max} & \rho \leq \rho_{crit} \\ c \frac{\ln \rho(x, t)}{\rho_{max}} & \rho > \rho_{crit} \end{cases}$$

Falls die kritische Dichte nicht erreicht ist, ist die Geschwindigkeit natürlich immer noch die maximal erlaubte Geschwindigkeit.

Frage: Warum ist diese Formel sinnvoll? beim letzten mal wurde für die Dichte eine andere Wahl getroffen.

Nun können wir den differentiellen Erhaltungssatz umschreiben, weswegen wir die Geschwindigkeit überhaupt berechnet haben.

Wir nehmen zunächst an, dass die Dichte linear ist und nur kleine Störungen auftreten. Dies führt dann zu der Erhaltungsformel:

$$\frac{\partial \rho_1}{\partial t} + q'(\rho_0) \frac{\partial \rho_1}{\partial x} = 0 \quad (2)$$

mit $\rho(x, t) = \rho_0 + \epsilon \rho_1(x, t)$ wobei ϵ die Störung ist.

Frage: Bedeutung der Formel?

Lineare DGL - gleichförmiger Verkehr

Wir wollen nun herausfinden, wie die Schwankungen/ Störungen ρ_1 aussehen. Also wollen wir eine Lösung der pDGL 2 finden. Dies können wir mithilfe der Methode der Charakteristik. Wir nehmen an, dass der Ort von der Zeit linear abhängt und wir wissen, dass gilt $x'(t) = c$. Also gilt: $x(t) = x_0 + tc$. Diese Geraden nennen wir Charakteristiken. ρ_1 bewegt sich konstant entlang dieser Charakteristik.

Frage: Warum?

Falls wir nun eine Anfangsbedingung gegeben haben mit $\rho_0(x(0), 0) = f(x_0)$, dann kann die Lsg von ρ_1 einfach berechnet werden: $\rho_1(x, t) = f(x - ct)$.

Dies muss ich noch üben!

Ausbreitung linearer Dichtewellen

Bei leichtem Verkehr, also wenn die Dichte kleiner ist als die kritische Dichte, bewegen sich Störungen mit positiver Geschwindigkeit.

Bei schwerem Verkehr jedoch, bewegen sich die Störungen mit negativer Geschwindigkeit, da die Autos, die in einer Region mit größerer Dichte sind, bremsen müssen. Das dahinter fahrende Fahrzeug ebenfalls. Diese Kettenreaktion wandert nun weiter.

Frage: Habe ich das genau verstanden?

Numerische Approximation

Nun wollen wir auch noch eine numerische Lsg der linearen DGL haben. Diese berechnen wir mittels Vorwärtsdifferenzschema für $\frac{\partial \rho}{\partial t}$ und Rückwärtsdifferenzschema für $\frac{\partial \rho}{\partial x}$ auf einem äquidistanten Gitter. Dabei kommt eine Formel heraus, die nur vom vorherigen Orts- und Zeitschritt abhängt.

Nun wollen wir den Bestimmtheitsbereich dieser Approximation bestimmen, also den Bereich auf der x-Achse (Ortsabhängig), vom dem die DGL abhängt.

Wenn der Bestimmtheitsbereich und der analytische Abhängigkeitsbereich, der Punkt auf der x-Achse, von dem der zu Approximierende Pkt via Charakteristik abhängt), nicht übereinstimmt, ist unsere Approximation bestimmt nicht korrekt, da der Punkt, von dem der zu approximierende Pkt wirklich abhängt, nicht in der Approximation betrachtet wurde.

Also wollen wir, dass der analytische Abhängigkeitsbereich im Bestimmtheitsbereich liegt. Diese Bedingung nennt sich Courant-Friedrich Lewis Bedingung und lässt sich als eine Bedingung an der Zeitschrittweite formulieren: Diese muss kleiner sein als die Ortsschrittweite geteilt durch die Steigung der Charakteristik.

Ungleichförmiger Verkehr

versteh ich noch gar nicht... wenn zeit in ruhe anschauen.

Anfahrvorgang an einer grünen Ampel

Was will ich hier modellieren? Was soll genau mein Ergebnis sein? Will ich die DGL lösen für dieses ρ mit meiner Anfangsverteilung? Aber was ist dann das Ergebnis?

Unstetige Verkehrsdichte

5 Molekulare Simulation