Dynamische Optimierung und Variationsrechnung

Timm Grams, Fulda, 15.11.2010 (rev. 06.02.2015)

Inhalt

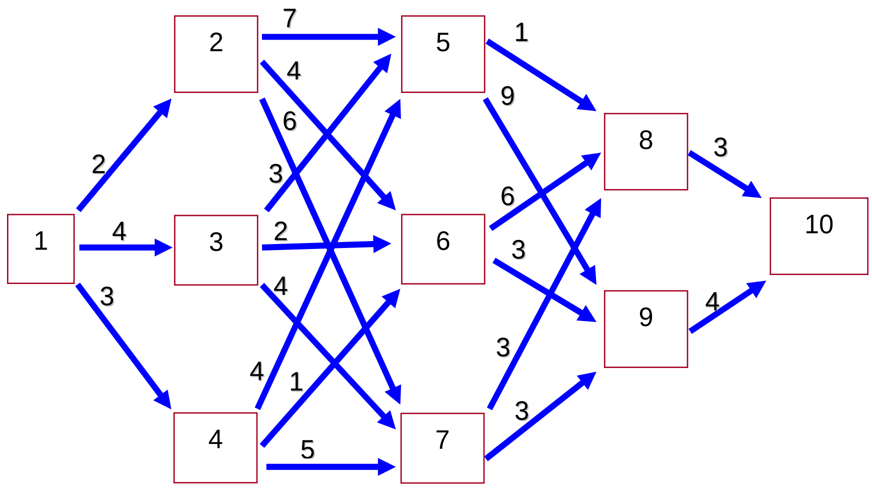
Einführendes Beispiel

Der abgebildete Digraph (gerichteter Graph) mit nichtnegativer Kantenbewertung hat zehn durchnummerierte Knoten. Es ist ein Weg längs der gerichteten Kanten zu finden, der vom Startknoten 1 zum Zielknoten 10 führt und dessen Kosten minimal sind. Die Kosten eines Weges ergeben sich als Summe der Bewertungen der zu durchlaufenden gerichteten Kanten. Beispielsweise ist der Weg 1-2-6-9-10 mit den Kosten 2+4+3+4 = 13 verbunden.

Grundsätzlich handelt es sich um ein *Problem der kürzesten Pfade*. Die Besonderheit dieses Netzes ist aber, dass es eine *Stufenstruktur* besitzt.

Der Zustand 0 ist der einzige Zustand der nullten Stufe. Die Zustände 2, 3 und 4 gehören zur ersten, die Zustände 5, 6 und 7 zur zweiten, die Zustände 8 und 9 zur dritten und der Zustand 10 zur vierten Stufe. Jede gerichtete Kante führt zu einem Zustand der unmittelbar darauf folgenden Stufe.

Diese Aufgabe lässt sich mit dem *Algorithmus von Dijkstra* lösen. Dabei bleibt die Stufenstruktur des Problems unberücksichtigt. Der Algorithmus der *dynamischen Optimierung* (Dynamic Programming) behebt diesen Mangel. Er nutzt die Stufenstruktur aus.



Das Einbettungsprinzip

Das Problem der kürzesten Pfade ist ein Musterbeispiel für die Anwendung von Problemlösungsheuristiken. Mit der *vollständigen Enumeration*, also dem vollständigen Auflisten aller möglichen Pfade und dem direkten Vergleich der jeweils zugehörigen Wegekosten, kommt man hier nicht weit. Die Anzahl der Möglichkeiten ist schon bei diesem einfachen Beispiel ziemlich groß.

Hier hilft uns die Heuristik der *Verallgemeinerung*. Wir formulieren diese Heuristik als *Einbettungsprinzip*: Sei $P(\alpha)$ ein Problem, das für eine bestimmte konstante Größe α formuliert ist. Im Beispiel ist α gleich dem Knoten 1. Die Größe α möge einer Menge D angehören, für deren Elemente sich das Problem in derselben Weise stellen lässt, D ist in unserem Fall die Menge sämtlicher Knoten 1, 2, 3, …, 10. Anstatt das Problem $P(\alpha)$ direkt anzupacken, betten wir es in die allgemeine Problemstellung $P(x)$ für alle $x \in D$ ein. Die Lösungen für diese Probleme beschreiben wir mit zunächst noch unbestimmten und vom Problem x abhängenden Größen. Dann werden Beziehungen zwischen diesen Lösungsansätzen gesucht. Unter Ausnutzung dieser Beziehungen beginnt man mit der Lösung der leichteren Probleme und geht dann weiter zu den schwerer lösbaren. So findet man dann möglicherweise eine allgemeine Lösung für die $P(x)$. Damit wäre dann auch das ursprüngliche Problem $P(\alpha)$ gelöst.

Das Standardproblem der dynamischen Optimierung

Das Standardproblem der dynamischen Optimierung wird für eine (endliche) Menge von möglichen Zuständen X definiert. Diese Zustände teilen sich auf mehrere aufeinander folgende Stufen auf: X_i, $i = 0, 1, 2, \ldots, N$. Es gilt also $X_i \cap X_j = \Phi$ für $i \neq j$ und $X = \bigcup_{i=0}^{N} X_i$. Für jeden Zustand x der Stufen 0 bis N-1 ist eine Menge von Folgezuständen $U(x)$ definiert: Wenn $x \in X_i$, dann ist $U(x) \subset X_{i+1}$. Eine Zustandfolge $(x_i, x_{i+1}, \ldots, x_j)$ heißt *zulässig*, wenn die Übergänge zwischen den Zuständen möglich sind, das heißt, wenn $x_{k+1} \in U(x_k)$ für $k = i, i+1, i+2, \ldots, j-1$. Solche Zustandsfolgen werden *Trajektorien* genannt. Für die Trajektorien und deren Abschnitte ist eine reellwertige nichtnegative Bewertungsfunktion g definiert: $0 \leq g(x_i, x_{i+1}, \ldots, x_j)$ für $i<j\leq N$. Die Aufgabe besteht darin, eine Trajektorie $(x_0, x_1, \ldots, x_N)$ vom Start X_0 ins Ziel X_N zu finden, deren Bewertung minimal ist.

Wir betrachten den Fall, dass sich der Wert einer Trajektorie aus den Werten der einzelnen Abschnitte additiv zusammensetzt; es gilt also $g(x_i, x_{i+1}, \ldots, x_j) = \sum_{k=i}^{j-1} g(x_k, x_{k+1})$.

Jetzt kommt das Einbettungsprinzip: Wenn es ausgehend von einem beliebig gewählten Zustand x überhaupt eine Trajektorie in die Zielmenge X_N gibt, dann ist darunter auch wenigstens eine optimale. Wir können jedem Zustand also einen Optimalwert S zuordnen: $S(x)$ ist gleich ∞, falls es keine zielführende Trajektorie gibt, ansonsten ist $S(x)$ gleich dem Wert einer optimalen Trajektorie.

Dynamische Optimierung: Der Algorithmus

Die bellmansche Funktionalgleichung

Die *Optimalwertfunktion S* genügt bei additiver Bewertungsfunktion der *bellmansche Funktionalgleichung:*

$$S(x) = \min_{y \in U(x)} g(x, y) + S(y) .$$

Beweis: Wir betrachten von x ausgehende Trajektorien $(x, y, \ldots, x_N)$, deren Endstücke $(y, \ldots, x_N)$ optimal sind. Für die Endstücke gilt also $S(y) = g(y, \ldots, x_N)$. Aus der Minimaleigenschaft von $S(x)$ folgt

$$S(x) \le g(x, y, \ldots, x_N) = g(x, y) + g(y, \ldots, x_N) = g(x, y) + S(y).$$

Zur Begründung: Falls $S(y)$ endlich ist, gibt es auch eine optimale Trajektorie $(y, \ldots, x_N)$ mit diesem Wert, und diese Trajektorie ist optimales Endstück der Trajektorie $(x, y, \ldots, x_N)$. Damit greift die Ungleichung. Falls $S(y) = \infty$, ist die die Ungleichung sowieso erfüllt. Sei nun $(x, x', x'', \ldots, x_N)$ eine von x ausgehende optimale Trajektorie. Es ist $S(x) = g(x, x', x'', \ldots, x_N) = g(x, x') + g(x', x'' \ldots, x_N)$. Das Endstück $(x', x'', \ldots, x_N)$ muss für den Startzustand x' optimal sein, denn gäbe es eine günstigere, würde die Ersetzung des Endstücks durch diese günstigere Trajektorie auch einen besseren Wert für $S(x)$ liefern, entgegen der Voraussetzung. Das heißt, es ist $S(x') = g(x', x'' \ldots, x_N)$. Es ist also $S(x) = g(x, x') + S(x')$. Mit $y = x'$ haben wir einen Folgezustand von x, der in der bellmanschen Funktionalgleichung das Minimum liefert: $S(x) = g(x, y) + S(y)$.

Rückwärtsrechnung

Für alle Zielzustände $x \in X_N$ wird $S(x)= 0$ gesetzt. Kosten fallen keine an. Sind für die Zustände y der Stufe $0<i$ bereits alle Optimalwerte bekannt, dann erhält man die Optimalwerte der Zustände x der Stufe i-1 mithilfe der bellmanschen Funktionalgleichung.

Beispiel: Der Optimalwert des einzigen Zustands der letzten (4.) Stufe ist gleich null: $S(10) = 0$. Für die Zustände der vorletzten (3.) Stufe ergeben sich aus der bellmansche Funktionalgleichung die Werte $S(9) = 4$ und $S(8) = 3$. So weitergehend ergeben sich sämtliche Werte der Optimalwertfunktion: $S(7) = 6$, $S(6) = 7$, $S(5) = 4$, $S(4) = 8$, $S(3) = 7$, $S(2) = 11$, $S(1) = 11$.

Vorwärtsrechnung

Eine optimale Trajektorie ergibt sich nun so: Ausgehend vom Startzustand x sucht man den Folgezustand y, für den die Gleichung $S(x) = g(x, y) + S(y)$ gilt. Das ist dann der nächste (Start-)Zustand der Trajektorie. Man ersetzt nun x durch y und wiederholt das Ganze. Das tut man so lange, bis ein Zielzustand erreicht ist.

Beispiel: Anwendung der Vorwärtsrechnung auf das einführende Beispiel – die Knoten spielen hier die Rolle der Zustände – liefert die folgenden optimalen Trajektorien: 1-3-5-8-10, 1-4-5-8-10 und 1-4-6-9-10. Sie haben jeweils den optimalen Wert von 11.

Nutzung der Politikfunktion

Bei der Vorwärtsrechnung wird eine Berechnung durchgeführt, deren Ergebnis bereits bei der Rückwärtsrechnung quasi nebenbei anfällt: Bei der Ermittlung des Optimalwerts eines Zustands wird notgedrungen auch ein optimaler Folgezustand bekannt. Den muss man sich nur merken. Dazu führen wir die *Politikfunktion p* ein: Für jeden Zustand x ist $p(x)$ der optimale Folgezustand: $\min_{y \in U(x)} g(x, y) + S(y) = g(x, p(x)) + S(p(x))$. Falls es keine zielführende Trajektorie gibt, wird $p(x)$ auf null gesetzt.

Die optimale Trajektorie $(x_0, x_1, \ldots, x_N)$, die von einem beliebigen Startpunkt $x_0 \in X_0$ ins Ziel X_N führt, ergibt sich aus der Rekursionsbeziehung $x_{i+1} = p(x_i)$. Gibt es mehrere Startpunkte, ist derjenige mit dem niedrigsten Optimalwert auszuwählen.

Probleme mit undefiniertem Planungshorizont

Ein einfaches Rendezvous-Problem

Ein Fahrzeug der Masse m hat den Abstand a vom Zielpunkt und ist *in Ruhe.* Das Fahrzeug kann mit einer Schubkraft K, $-K$ oder 0 in Richtung auf das Ziel beschleunigt werden. Gesucht ist eine Steuerung derart, dass das Fahrzeug in kürzester Zeit das Ziel erreicht und dort verharrt.

Systemgleichungen

Das Fahrzeug bewegt sich auf einer Geraden. Der Ort des Fahrzeugs wird durch s gemessen: $s = 0$ ist der Startpunkt und $s = a$ der Zielpunkt. Die Geschwindigkeit bezeichnen wir hier einmal mit r. Mit der (variablen) Schubkraft k ergeben sich die dynamischen Gleichungen

$$\dot{s} = r \tag{1a}$$

$$\dot{r} = k / m \tag{1b}$$

Das Beispiel dient der Demonstration von Optimierungsverfahren. Genaue Angaben von Masse, Kraft usw. sind hier ohne Belang.

Die Variable p möge für den Quotienten k/m stehen. Außerdem setzen wir den Quotienten K/m auf den Wert 1. (Denken Sie sich – je nach Lust und Laune – m/s^2, km/h^2 oder km/hs dazu.)

Die zweite der Systemgleichungen nimmt damit die folgende Gestalt an

$$\dot{r} = p \quad \text{mit } p \in \{-1, 0, 1\} \tag{1b'}$$

Diskretisierung

Die Zeitachse teilen wir in Abschnitte der gleichen Länge $\Delta t = 1$. Wir betrachten die physikalischen Größen Ort und Geschwindigkeit nur zu den Zeiten $t_i = i \cdot \Delta t = i$ ($i = 0, 1, 2, \ldots$). Die Schubkraft wird jeweils zu Beginn eines Intervalls festgelegt. Sie kann sich immer nur an den Intervallgrenzen ändern.

Das Euler-Cauchy-Verfahren liefert die folgenden *Systemgleichungen* für die diskretisierten Größen $s_i = s(t_i)$, $r_i = r(t_i)$ und $p_i = p(t_i)$:

$$s_{i+1} = s_i + r_i \tag{2a}$$

$$r_{i+1} = r_i + p_i \tag{2b}$$

Mit diesen diskretisierten Größen lässt sich das Rendezvous-Problem so formulieren: Gesucht ist eine möglichst kleine Zahl N und eine Folge von Steuerungsgrößen $(p_0, p_1, p_2,\dots ,p_{N-1})$ mit $p_i \in \{-1\ 0,\ 1\}$ derart, dass die sich aus den Systemgleichungen ergebenden Ortsfolgen $(s_0, s_1, s_2,\dots ,s_N)$ und Geschwindigkeitsfolgen $(r_0, r_1, r_2,\dots ,r_N)$ die *Randbedingungen*

$$s_0 = 0,$$

$$s_N = a,$$

$$r_0 = r_N = 0$$

erfüllen.

Da hier die Zahl N der Schritte gesucht ist, handelt es sich um ein Problem mit gesuchtem, also zunächst unbestimmtem, Planungshorizont.

Achtung: Hier wird das Euler-Cauchy-Verfahren sehr grobschlächtig angewendet. Das diskretisierte Problem ist dementsprechend nur eine sehr ungenaue Annäherung an das kontinuierliche! Aber hier geht es *nicht* um eine möglichst gute Beschreibung der Realität, sondern um die Erläuterung eines Optimierungsprinzips.

Problemstellung mit unbestimmtem Planungshorizont

Wir wollen das Rendezvous-Problem in eine Form bringen derart, dass es der dynamischen Optimierung zugänglich wird. Wir definieren die Zustände des Systems folgendermaßen: $x = (s, r, i)$. Die ganze Zahl i steht für den Zeitschritt und sie ist gleichzeitig Stufennummer. Gesucht ist eine möglichst kleine Zahl N und eine Folge $(x_0, x_1, \dots, x_N)$ mit $x_i = (s_i, r_i, i)$, so dass die Systemgleichungen (2a) und (2b) sowie die Randbedingungen $s_0 = 0$, $s_N = a$ und $r_0 = r_N = 0$ erfüllt sind.

Diese Problemstellung lässt sich verallgemeinern. Wichtig ist die Aufteilung der Zustände in zwei Komponenten, also $x = (z, i)$, derart, dass die Übergangsbedingungen der ersten Komponente unabhängig von der Stufe i sind. Präziser ausgedrückt: Die Zustandsmengen gestatten die Darstellung $X_i = (Z, i)$ mit einer von i unabhängigen Zustandsmenge Z für die z-Komponente. Ebenso ist die Menge der Folgezustände eines jeden Zustands $z \in Z$ von der Stufe unabhängig: $U(x) = U(z, i) = (U(z), i+1)$. (Das Symbol U wird hier in zwei Bedeutungen verwendet. Welche gemeint ist, kann man dem Argument entnehmen.)

Zeitinvariante dynamische Systeme passen in dieses Schema. Bei unserem diskreten Rendezvous-Problem muss man nur $z = (s, r)$ setzen. Was die Zustandbeschränkungen Z angeht, wurde bisher nichts vorausgesetzt. Jedoch ist es vernünftig, weder die Strecke noch die Geschwindigkeit über alle Grenzen wachsen zu lassen. Man kann also beispielsweise die Menge Z durch die Vorgabe von Minimal- und Maximalwerten für Ort und Geschwindigkeit begrenzen. Die Funktion U der Folgezustände ist in unserem Beispiel folgendermaßen definiert:

$$U(z) = U(s, r) = \{(s+r, r-1),\ (s+r, r), (s+r, r+1)\}$$

Die Forderung der Zeitunabhängigkeit wird auch an die nichtnegative Bewertungsfunktion gestellt: $g(x, x') = g((z,i), (z',i+1))$ hängt nur von den Zustandskomponenten z und z' ab und nicht etwa von i: $g((z,i), (z',i+1)) = g(z, z')$. Der Strich kennzeichnet einen jeweils zulässigen Nachfolgezustand.

Die Bewertungsfunktion bei unserem Rendezvous-Problem ist $g(z, z')=1$ für jedes zulässige Paar (z, z').

Nun lässt sich das Problem als eine Variante des Standardproblems der dynamischen Optimierung formulieren. Da die Zustandbeschränkungen stufenunabhängig sind, und da die Stufenzahl N unbestimmt ist, wäre die ursprüngliche Aufgabe trivial: Man bräuchte nur $N = 0$ zu setzen und erzielte null Kosten. Die Sache wird erst interessant, wenn man einschränkende Anfangs- und Zielbedingungen vorgibt: $z_0 \in A$ und $z_N \in \Omega$. Wir verlangen also, dass Lösungstrajektorie in einem Zustand der *Startmenge* A (sprich Alpha) startet und in einem Zustand der Zielmenge Ω (Omega) endet.

Problem der dynamischen Optimierung mit unbestimmtem Planungshorizont: Es sind eine Zahl N sowie eine zulässige Trajektorie $(z_0, z_1, \ldots, z_N)$ zu finden derart, dass die Bewertung der Trajektorie minimal wird und dass die Anfangs- und Zielbedingungen erfüllt sind. Es wird also verlangt, dass die für die Zustandsfolge $(z_0, z_1, \ldots, z_N)$ die folgende Bedingungen erfüllt sind

$$z_i \in Z \qquad \text{für } i = 0, 1, \ldots, N$$

$$z_{i+1} \in U(z_i) \qquad \text{für } i = 0, 1, \ldots, N\text{-}1$$

$$z_0 \in A,\ z_N \in \Omega$$

$$g(z_0, z_1, \ldots, z_N) = \sum_{k=0}^{N-1} g(z_k, z_{k+1}) = \min$$

Bellmansche Funktionalgleichung bei unbestimmtem Planungshorizont

Wir lösen das Problem nicht für eine feste Startmenge A, sondern wir betten das Problem in eine verallgemeinerte Problemstellung ein: Wir lösen das Problem für jedes beliebige z. Allerdings setzen wir voraus, dass die Trajektorienlänge durch n beschränkt ist: $N \le n$.

Wenn wir die Lösung all dieser Probleme haben, dann ist für hinreichend großen n auch die Lösung unseres ursprünglichen Problems dabei.

Die Optimalwertfunktion für das derart beschränkte Problem bezeichnen wir mit v_n; $v_n(z)$ ist also der Wert einer optimalen und maximal n-schrittige Trajektorie, die vom Anfangszustand z ins Zielgebiet Ω führt.

Wir fangen mit $n = 0$ an. Die Zustände, von denen man in 0 Schritten ins Ziel kommt, sind genau die Zustände der Zielmenge Ω. Für diese Zustände kennen wir den Optimalwert. Er ist gleich 0. Für alle anderen Zustände wird der Wert auf unendlich gesetzt:

$$v_0(z) = 0 \quad \text{für } z \in \Omega$$
$$v_0(z) = \infty \quad \text{sonst}$$

Sei nun die Optimalwertfunktion v_{n-1} für die Zielerreichung in maximal n-1 Schritten bereits bestimmt, dann kommt man zur Optimalwertfunktion v_n für maximal n-schrittige Trajektorien über die *bellmansche Funktionalgleichung*:

$$v_n(z) = \min_{z' \in U(z)} g(z,z') + v_{n-1}(z')$$, falls dieses Minimum kleiner als $v_{n-1}(z)$ ist;

ansonsten wird $v_n(z) = v_{n-1}(z)$ gesetzt.

Durch Vergrößerung des n erhält man für jedes z den Optimalwert, falls es einen solchen gibt. Im Falle der Rendezvous-Problems lässt sich diese Aussage präzisieren: Gilt für einen Zustand z und für ein n, dass $v_n(z) < \infty$, so ist $v_n(z)$ der gesuchte Optimalwert für diesen Zustand und dieser Wert lässt sich durch Vergrößerung des n nicht verbessern.

Wir lösen unser einfaches Rendezvous-Problem für den Zielort $s = a = 10$. Damit ergibt sich die Optimalwertfunktion der Tabelle. Die farbigen Felder markieren eine der optimalen Zustandsfolgen.

Die Optimalwertfunktion

s→ r↓	0	1	2	3	4	5	6	7	8	9	10
5	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞
4	4	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞
3	5	4	4	4	3	∞	∞	∞	∞	∞	∞
2	5	5	4	4	4	3	3	2	∞	∞	∞
1	6	5	5	5	4	4	3	3	2	1	∞
0	7	6	6	6	5	5	4	4	3	2	0
-1	∞	8	7	7	7	6	6	5	5	4	3
-2	∞	∞	∞	9	8	8	7	7	7	6	5

Zum Vergleich: Das Bang-bang-Prinzip

Eine weitere Lösungsidee beruht auf dem *Bang-bang-Prinzip*: Man fährt anfangs Vollgas und dann macht man eine Vollbremsung so, dass das Fahrzeug im Zielpunkt zur Ruhe kommt. Also ist in Gleichung (1b') anfangs die rechte Seite gleich 1. Dann wir umgeschaltet auf maximale Bremskraft, also auf -1.

Die allgemeine Lösung der dynamischen Gleichungen sieht in diesen Fällen mit den reellen Parametern A und B so aus:

$$s = A + Bt \pm t^2/2$$

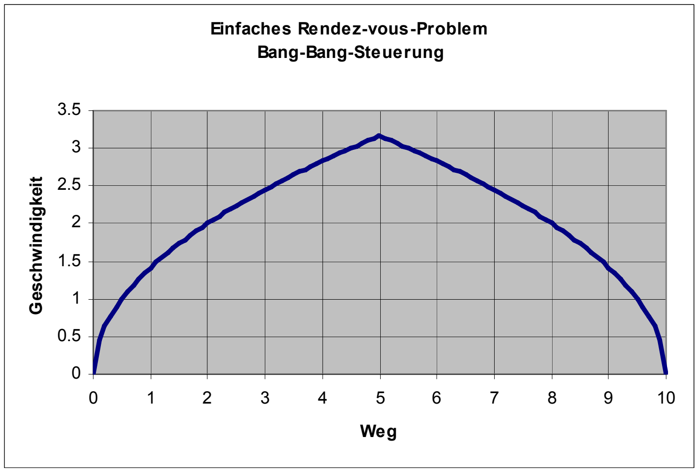
$$r = B \pm t$$

Für die erste Phase, das ist die Beschleunigungsphase und es gilt das positive Vorzeichen, sind die Randbedingungen $s(0) = r(0) = 0$ zu erfüllen: Anfangs sind Ort und Geschwindigkeit gleich null. Wir erhalten $r = t$ und $s = t^2/2$. Für das Weg-Geschwindigkeitsdiagramm drücken wir r in Abhängigkeit von s aus. Wir eliminieren die Zeit und erhalten für die Beschleunigungsphase die Formel $r = \sqrt{2s}$.

In der Bremsphase (negatives Vorzeichen) muss das Fahrzeug im Endpunkt a und zum (noch unbekannten) Endzeitpunkt T zur Ruhe kommen: $s(T) = a$, $r(T) = 0$. Wir erhalten die Formeln $r = T - t$ und $s = a - (T-t)^2/2$. Die Wegabhängigkeit der Geschwindigkeit ist dann gegeben durch $r = \sqrt{2(a-s)}$.

Der Moment des Umschaltens von Vollgas auf Vollbremsung ist genau dann, wenn die Geschwindigkeiten in den jeweiligen Phasen gleich sind: $\sqrt{2s} = \sqrt{2(a-s)}$. Das ist auf halbem Weg, also bei $s = a/2$, der Fall. Bis dahin ist die Zeit $\sqrt{a}$ vergangen. Am Anfang des Bremsbereichs ist $s = a/2 = a - (T-\sqrt{a})^2/2$. Daraus folgt, dass die Bremsphase genau so lange dauert wie die Beschleunigungsphase und dass die Gesamtzeit durch $T = 2\sqrt{a}$ gegeben ist. Das ist der optimale Wert. Im Beispiel wurde $a = 10$ gewählt. Es ergibt sich der Wert T=6.3. Der mittels dynamischer Optimierung gefunden Wert von 7 kommt an diesen Optimalwert schon ziemlich gut heran. (Mehr ist angesichts der groben Diskretisierung auch nicht zu erwarten.)

Die Tabelle der Optimalwertfunktion gibt – trotz der sehr groben Diskretisierung – einen ersten Eindruck von der Wegabhängigkeit der Geschwindigkeit. Zum Vergleich wird im Anschluss das Weg-Geschwindigkeitsdiagramm der Bang-bang-Steuerung dargestellt.



Die schnellste Bahn – ein Grundproblem der Variationsrechnung

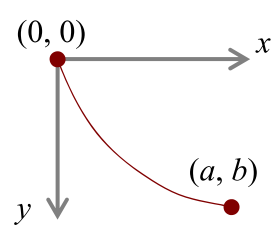
Johann Bernoulli (1667-1748) stellte im Jahr 1696 das berühmte Brachistochronenproblem (brachys heißt schnell und chronos ist die Zeit): „Zwei gegebene Punkte, welche verschiedenen Abstand vom Erdboden haben und nicht senkrecht übereinander liegen, sollen durch eine Kurve verbunden werden, auf welcher ein beweglicher Körper vom oberen Punkte ausgehend schnellstmöglich zum unteren Punkt gelangt.“ Nehmen Sie an, dass der Körper zu einem Punkt geschrumpft ist und reibungsfrei die Strecke durchläuft.

Zum Brachistochronenproblem gibt es einen eindrucksvollen Versuch im Mathematikum Gießen. Siehe dazu das Begleitbuch zur Ausstellung „Mathematik zum Anfassen“. Im Internet: www.Mathematikum.de. Johann Bernoullis Lösungsansatz, der auf einer Analogie zum snelliusschen Brechungsgesetz basiert, wird von Georg Pólya beschrieben (1988, S. 229 ff.).

Mit der Behandlung des Brachistochronenproblems begründete Johann Bernoulli das Gebiet der Variationsrechnung. Sein erster Lösungsansatz auf Basis des snelliusschen Brechungsgesetzes legt allerdings einen anderen Zugang nahe, den der dynamischen Optimierung. Und diesen wähle ich.

Ausgehend vom konkreten Problem werde ich schnell auf allgemeine Formulierungen wechseln, so dass eines der fundamentalen Ergebnisse der Variationsrechnung, die Euler-Lagrange-Differentialgleichung, in schöner Allgemeinheit erscheint. Was die Formelsprache angeht, halte ich mich weitgehend an das Buch von Dreyfus (1965).

Von der Bewegungsgleichung zur bellmanschen Funktionalgleichung



Das Koordinatensystem wird so festgelegt: Die x-Achse verläuft horizontal und die y-Achse ist zum Erdmittelpunkt, also nach unten gerichtet. Der Körper startet im Punkt (0, 0) und sein Zielpunkt ist der Punkt (a, b).

Im Startpunkt hat der Körper die Geschwindigkeit 0. Der Energieerhaltungssatz liefert für die Geschwindigkeit des Massepunkts in Richtung der Bahnkurve den Wert

$$v = \sqrt{2gy}\ .$$

Mit g wird die Erdbeschleunigung (Fallbeschleunigung) bezeichnet, die in unseren Breiten etwa 9.81 m/s^2 beträgt. Wir betrachten nun ein kleines Stück der durchlaufenen Strecke, ausgehend von (x, y) nach $(x+\Delta x, y+\Delta y)$. Wir nehmen das Stückchen als näherungsweise gerade an und bezeichnen die Steigung mit p:

$$\Delta y = p \cdot \Delta x.$$

Der Körper benötigt zum Durchlaufen dieser Teilstrecke der Länge

$$\Delta l = \sqrt{\Delta x^2 + \Delta y^2} = \Delta x \sqrt{1 + p^2}$$

näherungsweise die Zeit

$$\Delta t = \frac{\Delta l}{v} = \Delta x \cdot F(x, y, p) \quad \text{mit} \quad F(x, y, p) = \sqrt{\frac{1 + p^2}{2g \cdot y}}\ .$$

Im konkreten Fall hängt die Funktion F zwar nicht von der Variablen x ab Die Variable wird hier im Sinne einer Verallgemeinerung der Ergebnisse eingeführt.

Wir betten die konkrete Problemstellung, nämlich die Suche nach dem zeitoptimalen Weg, der vom Punkt (0, 0) zum Punkt (a, b) führt, in die allgemeine Problemstellung ein, die den Anfangspunkt (x, y) variabel hält. Für die Zusammenhänge zwischen den Problemen wollen wir dann wieder die bellmansche Funktionalgleichung formulieren. Wir müssen also zunächst dafür sorgen, dass die *bellmansche Optimalitätsbedingung,* die wir beim Beweis der Bellmanschen Funktionalgleichung oben benutzt haben, erfüllt ist: Teilwege, insbesondere Endstücke optimaler Wege sind selbst wieder optimale Wege. Diese Bedingung lässt sich erfüllen, wenn wir in der allgemeinen Problemstellung die Anfangsgeschwindigkeit nicht auf 0 sondern auf die von y abhängige Geschwindigkeit $v = \sqrt{2gy}$ setzen.

Die *Optimalwertfunktion* $S(x, y)$ ist gleich der Zeit, die der Körper ab dem Punkt (x, y) benötigt um zum Zielpunkt (a, b) zu kommen, also die Zeit für das Durchlaufen einer optimalen Kurve. Die *bellmansche Funktionalgleichung* nimmt die Form

$$S(x, y) = \min_{p} \big(\Delta x \cdot F(x, y, p) + S(x + \Delta x, y + p \cdot \Delta x) \big)$$

an. Sie besagt: Der Optimalwert $S(x, y)$, also die Zeitdauer für das Durchlaufen der optimalen Kurve von (x, y) nach (a, b), setzt sich zusammen aus der Zeit für das Durchlaufen des Streckenstücks von (x, y) nach $(x+\Delta x, y+p\cdot\Delta x)$ plus die Optimalzeit von dort zum Ziel. Die Steigung p des Streckenstücks ist genau so zu wählen, dass der Minimalwert erreicht wird.

Die Optimalwertfunktion S und die optimale Steigung p sind für alle Punkte (x, y) prinzipiell bestimmbar, zumindest im Diskreten. Für die Berechnung geht man vom Zielpunkt Schritt für Schritt um jeweils die Schrittweite Δx zurück und bestimmt für die zugehörigen y-Werte die Funktionen aus den bereits bekannten Werten. Dieser Idee folgend, erhält man auch eine einfache numerische Methode zur Lösung des Brachistochronenproblems durch Rasterung von x- und y-Achse.

Unter der Voraussetzung, dass die Differenzierbarkeitsbedingungen erfüllt sind, muss im Optimalpunkt die Ableitung des Klammerausdrucks nach p gleich null sein. Daraus ergibt sich die Bedingung

$$S_y = -F_p \,.$$

Die tief gestellten Variablenbezeichner besagen, dass nach der Variablen partiell abgeleitet wird. Die Optimalwerte für p legen die *Politikfunktion* $p = p(x, y)$ fest. Mit diesen Bezeichnungen lässt sich die Optimalitätsbedingung ausführlich so schreiben:

$$S_y(x,y) = -F_p(x,y,p(x,y))\,. \tag{1}$$

Für den Optimalwert wird in der bellmanschen Funktionalgleichung das Minimum erreicht und nach Einsetzen der Politikfunktion erhalten wir

$$S(x,y) = \Delta x \cdot F(x,y,p(x,y)) + S(x+\Delta x, y + p(x,y)\cdot \Delta x)\,.$$

Der Grenzübergang $\Delta x \to 0$ liefert eine zweite Gleichung

$$S_x(x,y) + S_y(x,y)\cdot p(x,y) = -F(x,y,p(x,y))\,. \tag{2}$$

Durch Einsetzen von (1) in (2) erhalten wir eine explizite Darstellung auch für S_x:

$$S_x(x,y) = F_p(x,y,p(x,y))\cdot p(x,y) - F(x,y,p(x,y))\,. \tag{2'}$$

Euler-Lagrange-Differentialgleichung und Beltrami-Identität

Die expliziten Darstellungen für S_x und S_y in den Gleichungen (1) und (2') eröffnet eine einfache Möglichkeit, die Optimalwertfunktion zu eliminieren. Wenn wir eine Gleichung für die Politikfunktion haben, lässt sich, ausgehend vom Startpunkt, die optimale Kurve hin zum Zielpunkt bestimmen, da dann ja in jedem Punkt (x, y) die Steigung dieser Kurve bekannt ist.

Ableitbarkeit vorausgesetzt, leiten wir nun die linke und rechte Seite von (1) nach x ab.

$$S_{yx}(x,y) = -F_{px}(x,y,p(x,y)) - F_{pp}(x,y,p(x,y))\cdot p_x(x,y)$$

Entsprechend leiten wir (2') partiell nach y ab.

$$S_{xy}(x,y) = F_{py}(x,y,p(x,y))\cdot p(x,y) + F_{pp}(x,y,p(x,y))\cdot p(x,y)\cdot p_y(x,y)$$
$$+ F_p(x,y,p(x,y))\cdot p_y(x,y) - F_y(x,y,p(x,y)) - F_p(x,y,p(x,y))\cdot p_y(x,y)\,.$$

Wegen $S_{xy} = S_{yx}$ ergibt sich daraus die Formel

$$F_{pp}\cdot p_x + F_{px} + F_{py}\cdot p + F_{pp}\cdot p\cdot p_y - F_y = 0\,.$$

Das ist die *Euler-Lagrange-Differentialgleichung,* die etwas kompakter geschrieben so aussieht:

$$\frac{\mathrm{d}F_p(x,y,p(x,y(x)))}{\mathrm{d}x} = F_y(x,y,p(x,y(x)))\,. \tag{3}$$

Dem Skript von Mertens (2008) entnehme ich die Anregung zur folgenden Rechnung:

$$\frac{\mathrm{d}}{\mathrm{d}x}(F - pF_p) = F_x + F_y\cdot p + F_p\cdot\frac{\mathrm{d}p}{x} - \frac{\mathrm{d}p}{x}F_p - p\frac{\mathrm{d}F_p}{\mathrm{d}x}$$

Der dritte und vierte Term heben sich auf. Wegen (3) ergeben auch der zweite und der letzte Term zusammen null. So erhalten wir

$$\frac{\mathrm{d}}{\mathrm{d}x}(F - pF_p) = F_x .$$

Beim Brachistochronenproblem hängt F nicht von x ab. Daraus folgt mit einer Konstanten c die *Beltrami-Identität:*

$$F - pF_p = c . \tag{4}$$

Lösung des Brachistochronenproblems

Im Fall der Brachistochrone ist $F(x, y, p) = \sqrt{\frac{1+p^2}{2g \cdot y}}$. Die partielle Ableitung nach p liefert

$$F_p = \frac{p}{\sqrt{2g \cdot y \cdot (1+p^2)}} .$$

Das in (4) eingesetzt ergibt:

$$\frac{1}{\sqrt{2g \cdot y \cdot (1+p^2)}} = c .$$

Wegen $p = \mathrm{d}y/\mathrm{d}x = y'$ erhält man eine gewöhnliche Differentialgleichung für die optimale Bahnkurve:

$$y \cdot (1 + y'^2) = \frac{1}{2g \cdot c^2} . \tag{5}$$

Jetzt erfolgt ein kühner Sprung. Wir setzen die Bahnkurve mit der noch wählbaren Konstanten k in einer Parameterdarstellung an: $x(\alpha) = k \cdot (\alpha - \sin \alpha)$ und $y(\alpha) = k \cdot (1 - \cos \alpha)$. Um zu diesem Ansatz zu kommen, haben die Großen der Mathematik sicherlich viele Wege und Ansätze ausprobiert. Wir wollen ihnen für den Lösungshinweis dankbar sein.

Da $y' = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}\alpha}{\mathrm{d}x/\mathrm{d}\alpha} = \frac{\sin\alpha}{1-\cos\alpha}$ folgt $y \cdot \left(1 + y'^2\right) = 2k$. Wir setzen für $2k$ den Wert der rechten Seite der Gleichung (5) ein. So haben wir die Lösung der Gleichung (5) und damit die des Brachistochronenproblems in Parameterdarstellung erhalten:

$$x = \frac{\alpha - \sin\alpha}{4g \cdot c^2}$$

$$y = \frac{1 - \cos\alpha}{4g \cdot c^2}$$

Es handelt sich um eine spitze Zykloide. Auf solchen Kurven bewegen sich die Peripheriepunkte eines abrollenden Rades mit dem Radius $1/(4gc^2)$. Der Abrollwinkel durchläuft die Werte von 0 bis A. Der finale Abrollwinkel A und die Konstante $4gc^2$ ergeben sich aus den Bedingungen $x(A) = a$ und $y(A) = b$. Die Hauptschwierigkeit ist die Bestimmung des finalen Abrollwinkels A. Die Endbedingungen $a = x(A)$ und $b = y(A)$ liefern die Gleichung

$$\frac{b}{a}=\frac{1-\cos A}{A-\sin A}.$$

Daraus lässt sich das passende A mit dem Newton-Verfahren iterativ ermitteln. Die Konstante c ergibt sich aus einer der Endbedingungen, beispielsweise aus $b = \mathrm{y}(A)$:

$$c=\sqrt{\frac{1-\cos A}{4gb}}.$$

Die Geschwindigkeit längs der Bahn ist gegeben durch

$$\frac{\mathrm{d}l}{\mathrm{d}t}=v=\sqrt{2gy}=\sqrt{\frac{1-\cos\alpha}{2c^2}}.$$

Die Ableitung der Streckenlänge nach dem Winkel α erhalten wir über die Beziehungen

$$\frac{\mathrm{d}x}{\mathrm{d}\alpha}=\frac{1-\cos\alpha}{4gc^2} \quad \text{und} \quad \frac{\mathrm{d}y}{\mathrm{d}\alpha}=\frac{\sin\alpha}{4gc^2}$$

folgendermaßen:

$$\frac{\mathrm{d}l}{\mathrm{d}\alpha}=\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\alpha}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\alpha}\right)^2}=\frac{1}{4gc^2}\cdot\sqrt{2(1-\cos\alpha)}.$$

Es gilt $v=\frac{\mathrm{d}l}{\mathrm{d}t}=\frac{\mathrm{d}l}{\mathrm{d}\alpha}\cdot\frac{\mathrm{d}\alpha}{\mathrm{d}t}$. Mit den eben berechneten Ableitungen führt das auf den Zusammenhang

$$\sqrt{\frac{1-\cos\alpha}{2c^2}}=\frac{1}{4gc^2}\cdot\sqrt{2(1-\cos\alpha)}\cdot\frac{\mathrm{d}\alpha}{\mathrm{d}t},$$

der sich zu

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t}=2gc$$

vereinfachen lässt. Zwischen dem Abrollwinkel und der Zeit besteht demnach ein linearer Zusammenhang

$$\alpha = 2gc\cdot t.$$

Damit ergibt sich die Gesamtzeit T für das Durchlaufen der Brachistochrone zu

$$T=A/2gc=A\cdot\frac{\sqrt{b/g}}{\sqrt{1-\cos A}}.$$

Die folgende Seite zeigt den Verlauf der Brachistochrone für den Radius $1/(4gc^2) = 1$.

Literatur

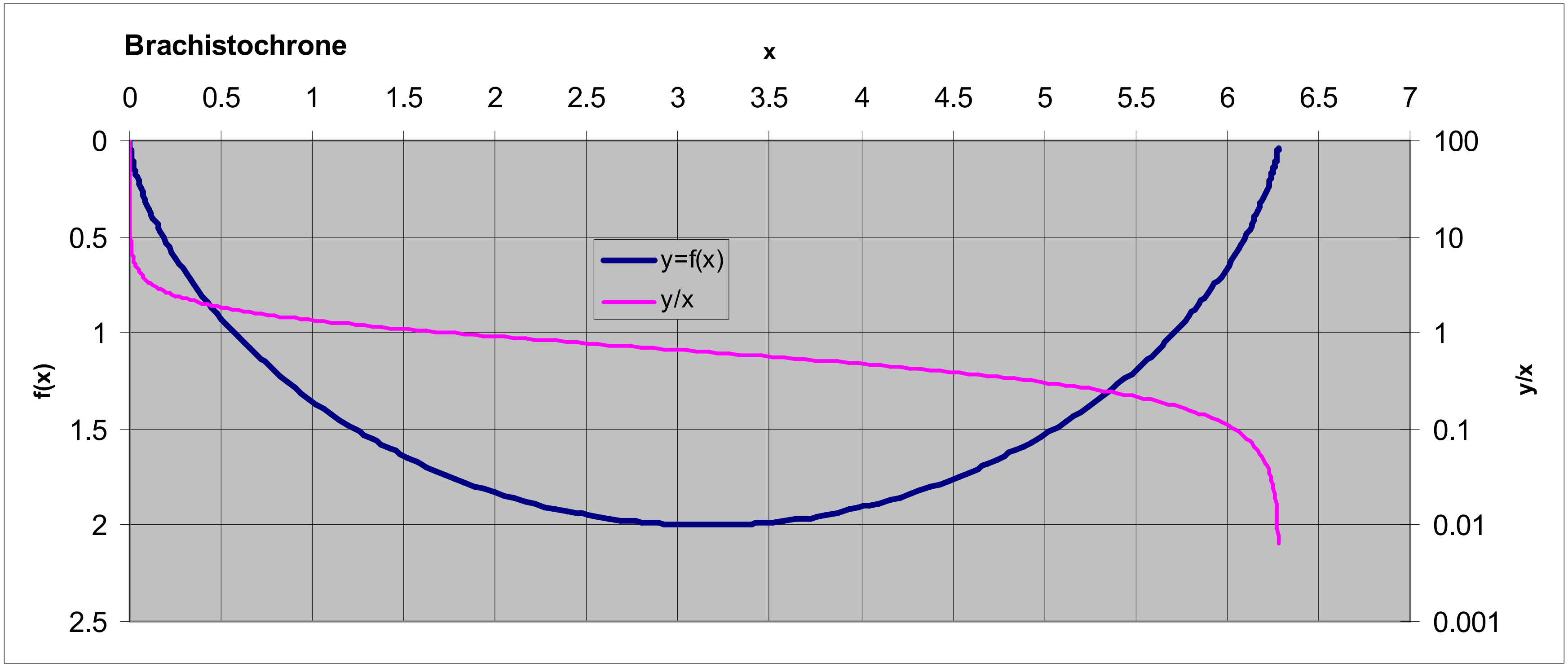
Dreyfus, Stuart E.: Dynamic Programming and the Calculus of Variations. Academic Press, New York 1965
Gaede, Karl-Walter; Heinhold, Josef: Grundzüge des Operations Research. Teil 1. Hanser, München, Wien 1976
Grams, Timm: Skriptum zu Informatik III. Fulda, 2002 ff.

Mertens, Stephan: Skriptum „10 Variationsrechnung“. Institut für Theoretische Physik, Otto-von-Guericke Universität, Magdeburg 2008

Neumann, Klaus: Dynamische Programmierung. BI, Mannheim 1969

Neumann, Klaus; Morlock, Martin: Operations Research. Hanser, München, Wien 1993

Pólya, Georg: Mathematik und plausibles Schließen. Band 1. Induktion und Analogie in der Mathematik. Basel 1988 (Originalausgabe: Mathematics and Plausible Reasoning, Vol. 1: Introduction and Analogy in Mathematics. Princeton 1954)


Brachistochrone
x
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
f(x)
0
0.5
1
1.5
2
2.5
y/x
100
10
1
0.1
0.01
0.001
y=f(x)
y/x