

§12 Schlussbemerkungen

Mit der Diskussion der kontinuierlichen Prozesse haben wir unsere Untersuchungen über dynamische Programmierung und Markov-Kette abgeschlossen. Wir haben gesehen, dass die Analysis der diskreten und kontinuierlichen Markov-Ketten sehr ähnlich ist. Im diskreten Fall ist die z -Transformation eine sehr nützliche analytische Methode, während im kontinuierlichen Fall die Laplace-Transformation diese Rolle übernimmt. In beiden Situationen erlaubt uns die passende Transformation, die Spezialfälle der Periodizität und mehrere ergodische Klasse zu analysieren, die so oft andere analytische Untersuchungen erschweren.

Selbst im bewerteten Prozess erweisen sich die Transformations-Methoden als nützlich für die Berechnung der total zu erwartenden Erlöse als Funktionen der Zeit und zur Bestimmung der asymptotischen Formen der Ausdrücke für diese Erlöse. Für ein System, das unter einer festen Politik läuft, verschafft uns die Kenntnis der total zu erwartenden Erlöse ein völliges Versehen des Systems.

Sehr interessant gestaltet sich das Problem, wenn für die Handhabung des Systems verschiedene Strategien zur Verfügung stehen. Im allgemeinen versucht man eine Menge von Strategien oder eine Politik zu bestimmen, welche zum Maximum des total zu erwartenden Erlöses führt. Wenn wir bei einem diskreten System den total zu erwartenden Erlös maximieren wollen, und zwar nur über wenige Stufen des Prozesses, ist eine Wertiterationmethode angezeigt. Jedoch in einem Prozess mit unbestimmter Dauer ist die Politik-Iterationsmethode vorteilhafter. Diese Methode bestimmt die Politik, die einen höheren durchschnittlichen Erlös pro Übergang als irgendeine andere in Betracht fallende Politik aufweist. sogar in Prozessen, in denen mehrere ergodische Klassen möglich sind, entstehen keine Schwierigkeiten von Bedeutung. Die damit verbundenen rechnerischen Arbeiten sind einfach, praktisch und leicht anzuführen.

Wenn wir nun aber den total zu erwartenden Erlös für ein kontinuierliches System maximieren, so ist unsere Wahl beschränkt. Das kontinuierliche Analogon der Wertiterationmethode ist so mühsam, dass uns die Praxis zwingt, Vereinfachungen vorzunehmen. Wenn wir speziell an Prozessen von kurzer Dauer interessiert sind, so ist der einfachste Weg, den kontinuierlichen Prozess durch einen diskreten Prozess zu approximieren und dann die Wertiteration anzuwenden. andererseits kann man bei Prozessen von langer Dauer die Politik-Iterationsmethode ebenso gut anwenden wie im diskreten Fall. Ferner sind die rechnerischen Anforderungen der beiden Prozessstypen so ähnlich, dass dasselbe allgemeine Computerprogramm für die Lösung beider Problemgattungen genügt. Abschliessend können wir sagen, dass die Politik-Iterationsmethode für die Lösung der kontinuierlichen Prozess ganz besonders wichtig ist, da praktische Alternativen fehlen.

Wir haben festgestellt, dass die Diskontierung den Grundcharakter des Entscheidungsproblems nicht ändert. Die früheren Feststellungen beim Vergleich der Wert- und Politik-Iterationsmethoden für diskrete und kontinuierliche Prozesse gelten ebenfalls für die Probleme mit Diskontierung. Die Diskontierung hat jedoch einige sehr interessante Merkmale. Erstens wird bei Prozessen von langer Dauer der Begriff des Gewinns durch denjenigen des Gegenwartswertes ersetzt, und unser Ziel bei der Politik-Verbesserung ist, die Gegenwartswerte aller Zustände zu maximieren. Zweitens können wir in unseren Berechnungen die Klassenstruktur des Prozesses ignorieren. Drittens bestehen Wertebereiche des Diskontierungsfaktors, die für jeden ihrer Werte die gleiche optimale Politik haben. Diese Eigenschaften ändern jedoch das rechnerische Verfahren nur wenig. Ein gut entworfenes Computerprogramm kann diskrete und kontinuierliche Prozesse, mit oder ohne Diskontierung, lösen. Wenn die Politik-Iterationsmethode angewendet wird, dann haben wir als Nebenprodukt der Berechnung der optimalen Politik eine Menge von Zustandswerten, die bei besonderen Umständen eine Auswertung des Abweichens von dieser Politik erlaubt. In den meisten Systemen sind diese Werte interessanter und nützlicher, als ursprünglich angenommen werden könnte. Es ist sehr wichtig, dass, wenn diese Zahlen gebraucht werden, wir daran denken, dass ihre Gültigkeit sich auf die Annahme stützt, dass die optimale Politik fast immer befolgt wird.

Die hier gegebenen Beispiele des Taxifahrers, des Automobilersatzproblems, etc. wurden so einfach dargestellt, dass sie nur eine Anregung für weitere Anwendungen darstellen können. Die Erwägungen über die Wahl der möglichen Anwendungen sind die folgenden. Erstens: Kann das System adäquat durch so wenig Zustände beschrieben werden, dass die Lösung der entsprechenden Gleichungssysteme rechnerisch möglich ist? Zweitens:

Sind die Daten, die notwendig sind, um die Strategien des Systems zu beschreiben, verfügbar? Sind die Antworten zu diesen Fragen positiv, haben wir eine mögliche Anwendung entdeckt. Wir haben allen Grund zu glauben, dass eine mögliche Anwendung, verbunden mit sorgfältiger Arbeit, eine erfolgreiche Anwendung liefert.