TITULO: Examples in Markov Decision Processes

AUTOR: A. B. Piunovskiy

EDITORIAL: London: Imperial College Press Optimization Series: Vol. 2.

SERIES ON OPTIMIZATION SERIES: VOL 2.

DISTRIBUTION: World Scientific Publishing Co. Pte. Ltd.

AÑO: 2013.

ISBN 978-1-84816-793-3 (Cubierta); ISBN 978-1908979-66-7(e-libro); 978184816794-0 (e-libro-sólo instituciones); ISSN 2041-1677

Precio: De 41€ a 50€.

URL: www.icpres.co.uk

MSC: Principal: 90 Operations research, mathematical programming. Clasificación principal: 90C40, 90C39, 6OJXX; Secundarios: 93E20, 49L20.

ISSN 2041-1677

Revisor: Francisco José Cano Sevilla. Profesor Universidad Complutense

RESUMEN

Este excelente libro contiene aproximadamente 100 ejemplos, que ilustran la teoría de los procesos de Markov, de tiempo discreto, controlados. La atención principal se presta a las propiedades inesperadas para el intuitivo contaje de los procesos de optimización. Tales ejemplos ilustran las condiciones impuestas a los teoremas conocidos sobre procesos de decisión markovianos. La pretensión era coleccionarlos juntos en un texto de referencia que pudiese ser considerado como un complemento a las monografías existentes sobre procesos de decisión markovianos. El libro es auto-contenido y unificado en la presentación.

Las exposiciones teóricas principales y las construcciones están dadas, y los ejemplos particulares pueden leerse unos independientemente de los otros. El texto Ejemplos en Problemas de Decisión Markovianos, es una fuente esencial de referencia para los matemáticos y para todos aquellos que aplican la teoría del control óptimo en cuestiones prácticas. Este libro recopila conjuntamente ejemplos basados sobre tales fuentes, junto con otras fuentes nuevas. En suma, indica las áreas en las que los procesos de decisión markovianos pueden usarse. Los investigadores activos pueden referirse a este libro como una aplicación de los teoremas y métodos matemáticos.

Los ejemplos presentados tratan bien de mantener soluciones de contaje intuitivo o bien ilustran la importancia de las condiciones en los teoremas conocidos. No todos los ejemplos son igualmente simples o complicados. Algunos ejemplos van dirigidos a estudiantes no graduados, mientras otros serían de interés a investigadores amateurs o profesionales.

Este libro consta de cuatro capítulos en línea con los cuatro tipos diferentes principales de los procesos de decisión markovianos; los modelos de horizonte finito, horizonte infinito con pérdida total o descontada, y pérdida media sobre un intervalo de tiempo infinito. Tales supuestos teóricos básicos y las demostraciones de las afirmaciones auxiliares se incluyen en los dos apéndices.

Se consideran sólo problemas de minimización. Cuando se formulan los teoremas y los ejemplos se publican en libros, revistas, artículos o documentos dedicados a la maximización, se ajustan los supuestos para nuestro caso sin ninguna observación especial. Debería enfatizarse que la terminología en los procesos de decisión markovianos no está completamente fijada.

Los procesos de decisión markovianos es una rama de las matemáticas basada en la teoría de la probabilidad, control óptimo y análisis matemático. Algunos libros sobre la materia con contraejemplos/paradojas en probabilidad existen en la literatura, y entonces no sorprende que los procesos de decisión markovianos estén completos, con ejemplos inesperados de contaje intuitivo.

La pretensión global de este atrayente texto es reunir juntamente tales ejemplos. ¡Fascinante tarea! Este libro debería ser considerado como un complemento a los textos científicos y monografías sobre procesos de decisión markovianos. El libro puede servir como un libro de referencia al que se puede volver para contestar a las curiosidades que surgen cuando se estudia o enseña procesos de decisión markovianos. Todos los ejemplos son auto-contenidos y pueden leerse independientemente unos de otros. En lo relativo a cadenas de Markov no controladas, debemos mencionar la brillante colección de ejemplos de Suhov y Kelber (2008).

Este libro es altamente recomendado. Algunos ejemplos están dirigidos a estudiantes no graduados, mientras otros serían de interés a avanzados no graduado, estudiantes tanto graduados como investigadores en teoría de la probabilidad, control óptimo y matemática aplicada, con objeto de tener una mejor comprensión de la teoría; expertos investigadores profesionales o amateurs en procesos de decisión markovianos.

**Capítulo 1: Modelos de horizonte finito (páginas 1-50)**

En este capítulo se analizan la definición y lo básico de los modelos de horizonte finito. La pretensión es encontrar una estrategia de control óptimo, resolviendo el problema relativo al funcional factible usando la aproximación de la programación dinámica, debido a que el sistema probabilístico evoluciona con el tiempo. Si las funciones de pérdida son no acotadas, la situación respecto a la aproximación por programación dinámica puede ser más complicada. El lema 1.1 y los corolarios 1.1 y 1.2 pueden ser útiles en la solución, debido a que el lema y los corolarios dan las condiciones suficientes de optimación. Los ejemplos presentados, en este capítulo, son 16 problemas interesantes, que permiten una ilustración completa y extensa de los modelos de horizonte finitos. Todos los ejemplos considerados son auto-contenidos.

Los ejemplos son los siguientes:

1. No transitividad de la correlación. La propiedad de poseer correlación positiva no es necesariamente transitiva. Esta cuestión está estudiada ampliamente en la literatura.
2. El control más frecuente no es el mejor. El contraejemplo propuesto es claro.
3. Votación. Este es un ejemplo interesante relacionado con la decisión final de acuerdo con la mayoría entre las tres opiniones de tres magistrados que investigan una persona acusada que es actualmente culpable.
4. El problema de la secretaria. Aquí se considera sólo una versión muy simple del problema clásico y famoso, estudiado con profundidad en la literatura especializada.
5. Optimación restringida. Este atractivo ejemplo muestra que el principio de Bellman no se verifica en la solución y la estrategia de control óptimo puede parecer extraña. Observar que, con frecuencia, la solución a un problema restringido de un proceso de decisión markoviano, viene dada por una estrategia de control aleatorizada de Markov; además, no existe ninguna razón para considerar estrategias dependientes del pasado. Introduciendo variables aleatorias artificiales, se debería modificar la pérdida total, y en este nuevo modelo, entonces el principio de Bellman se verifica.
6. Selectores de equivalentes markovianos en procesos de decisión markovianos no atómicos. En esta situación se presentan dos ejemplos diferentes que muestran que, en el caso no atómico, un selector φ no existe.
7. Selectores fuertemente equivalentes markovianos en procesos de decisión markovianos no atómicos. La noción de estrategias π and φ, fuertemente equivalentes es importante en la teoría del transporte de masas. El autor muestra en el teorema 1.1 asociado a este ejemplo, que todas las condiciones establecidas son importantes para la equivalencia fuerte.
8. Intercambio de stock. Merece la pena prestar atención a las diferentes funciones de pérdida empleadas, debido a que causan curiosas conclusiones en los problemas reales.
9. ¿Estrategia markoviana o no? ¿Aleatorización o no? ¿Cuando se viola el principio de Bellman? El ejemplo muestra que el requerimiento relativo a los infinitos es esencial. En el ejemplo propuesto, la estrategia de control φ no es uniformemente óptima, aún siendo correcto el razonamiento para un selector arbitrario, significa que las estrategias no aleatorizadas, no pueden satisfacer todas las igualdades establecidas y no pueden ser uniformemente óptimas.
10. Uniformemente óptima, pero no estrategia óptima. El ejemplo es el anterior, ligeramente modificado, ignorando la etapa inicial y colocando otra distribución de probabilidad.
11. Martingalas y el principio de Bellman. Si las funciones C y c, las funciones de pérdida están acotadas inferiormente, entonces el proceso de estimación es una martingala si y sólo si π es óptima. El ejemplo 9 y el propuesto aquí, muestran las dificultades de la afirmación anterior, debido a que el proceso estimador no es una martingala.
12. Convenios sobre la esperanza matemática y los infinitos. Estos ejemplos interesantes se dedican al uso de otras convenciones sobre esperanzas e infinitos, como algunos autores sugieren para calcular el criterio de realización, aceptando la regla (+∞) + (-∞) = +∞.
13. Función diferenciable en ninguna parte vt (x); función discontinua vt (x). En este ejemplo se muestra como una proposición en análisis. Es conocido que “una serie funcional puede converger (absolutamente) a una función continua, pero diferenciable en ninguna parte”.
14. La función de Bellman no medible. La aproximación a la programación dinámica se basa en la hipótesis de que la ecuación de Bellman tiene una solución medible. El siguiente ejemplo, muestra que la función de Bellman puede no ser medible Borel, aún en el caso más simple de ser T=1, C(x) ≡0, con una función de pérdida medible.
15. Ninguna estrategia es uniformemente ε-óptima. Este ejemplo aquí planteado es un antiguo ejemplo considerado en la literatura.
16. Modelo semi-continuo. Un proceso de decisión markoviano se dice semi-continuo si la condición 1.1 se satisface: La condición 1.1. establece 1) el espacio de acciones es compacto, 2) las probabilidades de transición es un núcleo estocástico continuo y 3) las funciones de pérdida son semi-continuas inferiores y acotadas inferiormente. Si el espacio de acciones no es compacto o la probabilidad de transición no es continua, o las funciones de pérdida no son semi-continuas inferiores, entonces algunos ejemplos triviales prueban que el selector deseado φ\* puede no existir.

**Capítulo 2:** **Modelos homogéneos de horizonte infinito: Pérdida total esperada (páginas 51-126)**

En este capítulo se supone que el horizonte temporal no es finito. Las definiciones de las estrategias y selectores son idénticas al caso de horizonte finito. El objetivo es encontrar una estrategia de control óptimo resolviendo el funcional realizable. Las siguientes condiciones se verifican siempre: “para cualquier estrategia de control, los valores esperados (positivos o negativos) son finitos” y también la condición de Putterman (194), Sección 7.2 sobre los modelos positivos.

Los procesos de decisión markovianos se dicen absorbentes si existe un estado, para el que el proceso controlado se absorbe en el tiempo T. Modelos absorbentes se consideran en los ejemplos2,7,10,13,16,17,19,20,21,24,28. Los ejemplos 3, 4, 9, 13,18 pertenecen al área de parada óptima en la que, en cada etapa, existe la posibilidad o bien poner el proceso de control en un estado especial absorbente, con ninguna pérdida futura.

Los modelos homogéneos de horizonte infinito con el criterio basado en la pérdida esperada total. En este capítulo se analizan 28 ejemplos agradables con el criterio anterior.

Los ejemplos son los siguientes:

1. Estrategias mixtas. Definiciones y ejemplo casi- trivial.
2. Soluciones múltiples a la ecuación de optimación. Aplicación a un modelo de colas de tiempo-discreto.
3. Modelo finito: soluciones múltiples a la ecuación de optimación; estrategia conservadora pero no igualadora. La condición de Putterman, se satisface, por lo que la función de Bellman coincide con la solución máxima no positiva.
4. La única estrategia conservadora no es igualadora y no optima. El ejemplo muestra que no existen ningunas estrategias óptimas en este modelo.
5. Cuando la estrategia de iteración no es buena. Si el modelo es positivo. y la función de coste está acotada, entonces el teorema debido a Strauch (1966) se verifica. Para modelos negativos, el teorema debido a Putterman (1994), dice que si la estrategia es tal que el algoritmo de iteración termina, entonces ésta es a su vez, una estrategia óptima.
6. Cuando el valor iterado no es bueno. Se propone un ejemplo con la violación de la primera condición.
7. Cuando el valor iterado no es bueno; modelo positivo I. El ejemplo propuesto: valor iterado no converge a la función de Bellman.
8. Cuando el valor iterado no es bueno: modelo positivo II. El ejemplo muestra que la pretensión, la existencia del límite, puede fallar si el modelo no es negativo.
9. Valor iterado y estabilidad en problemas de parada óptima.
10. Una estrategia no igualadora es uniformemente óptima. Ejemplo cuando la segunda condición se viola.
11. Un selector estacionario uniformemente ε-óptimo no existe (modelo positivo). El ejemplo propuesto lo demuestra.
12. Un selector estacionario uniformemente ε-óptimo no existe (modelo negativo). El ejemplo propuesto puede llamarse apuesta (juego).
13. Modelo negativo con acciones finitas, donde un selector estacionario uniformemente ε-óptimo no existe. Un ejemplo que puede ser reformulado como apostar.
14. Selectores óptimos uniformemente próximos. El ejemplo muestra que si el espacio de estados es contable el teorema de Ornstein (1969) se verifica. El ejemplo también muestra que el teorema no puede ser esencialmente mejorado.
15. Modelos semi-continuos y el dilema del chantajista. Gran número de resultados muy potentes son conocidos para los modelos semi-continuos. Se requiere una condición adicional: 1) el espacio de acciones es compacto, b) la función de pérdida para cada estado es semi-continua inferior en cada acción, y c) para cada estado, la función integral es continua en cada acción para toda función (medible) acotada u. El ejemplo es el dilema del chantajista (Bertsekas, 1987, pág. 254).
16. No un modelo semi-continuo. Si el modelo no es semi-continuo, entonces no se puede garantizar la existencia de estrategias óptimas. El modelo propuesto muestra que ningún selector es ε-óptimo.
17. La función de Bellman es no medible y ninguna estrategia es uniformemente ε-óptima
18. Una estrategia aleatorizada es mejor que cualquier selector (espacio de acciones finito)
19. La aproximación del fluido no trabaja.
20. La aproximación del fluido: modelo refinado.
21. Medidas de ocupación: soluciones fantasmas. Para una estrategia de control fijada, la medida de ocupación es una medida particular sobre XxA. La introducción de esta medida, hace que la aproximación analítica convexa a los procesos de decisión markovianos sea fructífera, especialmente en la optimación restringida. En lo que sigue se analizan diferentes ejemplos ilustrativos.
22. Medidas de ocupación en modelos transitorios
23. Medidas de ocupación y dualidad
24. Medidas de ocupación: compacticidad
25. La estrategia audaz (atrevida) en la apuesta no es óptima (límite de la banca o casa). This classical game can be modeled as an MDP.
26. La estrategia audaz (atrevida) en la apuesta no es óptima (inflación). Otro ejemplo del problema del apostador.
27. Estrategia de búsqueda para un objetivo móvil. En este ejemplo, parece plausible que la estrategia óptima sea muy simple, buscar aquel lugar que de la mayor probabilidad de encontrar el objeto.
28. El duelo a tres (truel). El truel secuencial es un juego que generaliza el duelo simple. Cada tirador desea maximizar su probabilidad de ganar el juego.

**Capítulo 3:** **Modelos homogéneos de horizonte infinito: Pérdida descontada (páginas 127-176)**

Este capítulo se dedica a problemas idénticos tratados con anterioridad donde β € (0, 1) es el factor descuento. Como es usual vπ es el funcional realizable. El modelo con descuento es un caso particular de los procesos de decisión markovianos con pérdida total esperada. El problema es ahora equivalente a la investigación del modelo modificado (absorbente), con tiempo de absorción espera totalmente acotado, finito.

Sin embargo, los modelos descontados tradicionalmente constituyen un área especial en los procesos de decisión markovianos. Por esta razón, los ejemplos siguientes, excepto algunos casos, son modificaciones especiales que introducen el factor descuento. 23 ejemplos diferentes, contiene este capítulo.

Los ejemplos son los siguientes:

1. Soluciones fantasmas de la ecuación de optimación.
2. Cuando el valor iterado no es bueno: modelo positivo.
3. Una estrategia no-óptima π tal que v π x resuelve la ecuación de optimación
4. La única estrategia conservadora no es igualadora y no óptima.
5. Valor iterado y convergencia de estrategias.
6. Valor iterado en modelos contables
7. La función de Bellman es no medible y ninguna estrategia es uniformemente ε-óptima
8. Ningún selector es uniformemente ε-óptimo
9. Estrategias miopes. Una estrategia estacionaria, uniformemente optima en el modelo de una etapa homogéneo, T=1 con pérdida terminal C(x) =0, se dice miope.
10. Controladores estables e inestables para sistemas lineales
11. Acciones óptimas incorrectas en el modelo con información parcial
12. Medidas de ocupación y estrategias estacionarias
13. Optimación restringida y el principio de Bellman
14. Optimación restringida y multiplicadores de Lagrange
15. Optimación restringida: soluciones múltiples
16. Pérdida descontada ponderada y (N,∞)- selectores estacionarios
17. Descuento no constante
18. La estrategia óptima próxima o cercana no es Blackwell-óptima.
19. Estrategias optima-Blackwell y pérdida de oportunidad
20. Estrategias optima-Blackwell y óptimas-n descuento
21. Ningunas estrategias óptimas-Blackwell (Maitra)
22. Estrategias óptimas cuando β→ 1 y el proceso de decisión markoviano con pérdida media I
23. Estrategias óptimas cuando β→ 1 y el proceso de decisión markoviano con pérdida media II

**Capítulo 4:** **Modelos homogéneos de horizonte infinito: Pérdida media y otros criterios (páginas 177-252)**

Bajo condiciones más generales, los problemas de optimación, con pérdida media (u otros criterios), del funcional realizable están bien definidos, por ejemplo, si la función de pérdida c está acotada inferiormente. En este contexto, tales estrategias se dicen AC-óptimas, es decir, óptimas de coste medio. Si el modelo es finito, entonces existe un selector estacionario AC-óptimo. La situación llega a ser más complicada si bien el espacio X o bien el espacio A no son finitos. La programación dinámica es muy adecuada para este problema.

Algunas estrategias Markov y semi-Markov se consideran en el contexto del criterio AC-óptimo...

En este capítulo cuatro, una elección interesante y fascinante se analiza. En particular 31 ejemplos se examinan con los anteriores criterios.

Los ejemplos son los siguientes:

1. ¿Por qué limsup?
2. Estrategias no canónicas AC-óptimas
3. Ternas canónicas y ecuaciones canónicas
4. Soluciones múltiples a las ecuaciones canónicas en modelos finitos
5. Ningunas estrategias AC-óptimas
6. Ecuaciones canónicas sin soluciones: espacio de acciones finito
7. Ningunas estrategias estacionarias AC-óptimas en un modelo de estados finito
8. Ningunas estrategias AC-óptimas en un modelo semi-continuo de estados finito
9. Modelos semi-continuos y la suficiencia de selectores estacionarios
10. Ningunas estrategias estacionarias AC-óptimas en un modelo unicadena con un espacio finito de acciones
11. Ningunas estrategias estacionarias AC- ε-óptimas en un modelo de acciones finito
12. Ningunas estrategias de Markov AC- ε-óptimas
13. Perturbación singular de un proceso de decisión markoviano
14. Estrategias Blackwell-óptimas y AC-óptimas
15. Estrategia iterada en un modelo unicadena.
16. Estrategia iterada unicadena en un modelo de comunicación finito.
17. Estrategia iterada en modelos semi-continuos
18. Cuando el valor iterado no es bueno
19. La aproximación de horizonte finito no funciona
20. La aproximación de la programación lineal a los modelos finitos
21. Programación lineal en modelos infinitos. La programación lineal prueba ser efectiva en los modelos finitos. En el caso general, esta aproximación se ha desarrollado en la literatura, pero bajos condiciones especiales.
22. Programas lineales y frecuencias esperadas en modelos finitos.
23. Optimación restringida
24. Estrategias óptimas, AC, sesgos, adelantamientos y costes de oportunidad: modelo periódico
25. Estrategias óptimas, AC and adelantamientos medios
26. Estrategias óptimas, Blackwell, sesgos, adelantamientos y AC
27. Estrategias óptimas cercanas o próximas y adelantamientos medios óptimos
28. Adelantamientos medios óptimos Fuertes, estrategias AC-óptimas y pérdidas de oportunidad mínimas
29. Estrategias óptimas de adelantamientos fuertes y fuertes\*.
30. Paradoja de Parrondo
31. Una estrategia de servicio optima en un sistema de colas.

**Epílogo**

Brevemente mencionar varias aplicaciones a la vida real de los procesos de decisión markovianos

* Control de un objeto movible. El objetivo puede ser por ejemplo, alcanzar el objetivo con una mínima energía esperada. El estado es la posición del objeto sujeto a las perturbaciones aleatorias y la acción corresponde a la potencia del motor.
* Control de los recursos acuáticos. La función considerada consiste en maximizar la utilidad esperada del consume del agua. El estado es la cantidad de agua en el depósito y las decisiones se refieren al uso del agua. Consumption-investment problems. The objective is to minimize the total expected consumption over the planning interval.
* Control de inventarios. El objetivo consiste en maximizar el beneficio total esperado de venta del producto.
* Fiabilidad. En este caso, minimizar la pérdida total esperada resultante de los fallos y del coste de mantenimiento.
* Matemáticas financieras. El estado es la riqueza actual a lo largo del vector stock de precios en un Mercado aleatorio. La acción representa la reestructuración del portfolio auto-financiado, El objetivo es la maximización de la utilidad esperada asociada con la riqueza final.
* Venta de un active. El objetivo es maximizar el beneficio total esperado.
* Apuestas. El objetivo es maximizar la probabilidad de alcanzar el objetivo. Ejemplos de apuestas se dan en el capítulo 2, ejemplos 14, 25 y 26.

Existen en la literatura otros ejemplos significativos, por ejemplo, control de calidad en una línea de producción; gestión de bosques; poblaciones controladas; participación en un concurso de preguntas y respuestas; organización de la enseñanza y exámenes; optimación de esfuerzos publicitarios; seguros, y así sucesivamente.

**Apéndice A: Espacios de Borel y otras cuestiones teóricas (páginas 257-266)**

A.1. Conceptos principales. Definiciones, teoremas: Tychonoff, Urysohn, etc., espacios compactos, cubo de Hilbert y así sucesivamente.

A.2. Medidas de Probabilidad sobre espacios de Borel: Definiciones, existencia de núcleos estocásticos medibles, conceptos sobre un espacio métrico, etc.

A.3. Funciones semi-continuas y selección medible: Definiciones básicas, propiedades de las funciones inferiores y superiores, semi-continuas y acotadas inferiormente sobre espacios metrizables, y metrizables separables

A.4. Teorema abeliano (tauberiano)

**Apéndice B: Demostraciones auxiliares de lemas y proposiciones enunciadas (págs. 267-280)**

Este apéndice se dedica a las demostraciones auxiliares de los lemas 2.1 y 3.2, y de las proposiciones 4.1, 4.2 y 4.4 enunciadas.

**Notación (págs. 281-282)**

**Lista de los enunciados principales (págs. 283-284)**

**Bibliografía (págs. 285-290).** Existen alrededor de 69 referencias interesantes.

**Indice (páginas 291-293)**

**CONCLUSION**

Este libro podría ser considerado como un complemento a los libros de texto y monografías sobre los procesos de decisión markovianos. El libro también puede servir como un libro de referencia para aquellos que deseen volver a las curiosidades que surgen en el estudio o enseñanza de los procesos de decisión markovianos. Todos los ejemplos son auto-contenidos y pueden leerse independientemente unos de otros. Cuando se estudian o se usan métodos matemáticos, el estudiante avanzado o investigador debe comprender lo que ocurre si alguna de las condiciones impuestas en los teoremas rigurosas no se satisfacen. Gran número de ejemplos confirman la importancia de tales condiciones que fueron publicadas en artículos de revistas que a menudo son difíciles de encontrar.

En mi opinión, este notable y atrayente libro es altamente recomendado. Algunos ejemplos están dirigidos a estudiantes no graduados, mientras otros son de interés a estudiantes no graduados avanzados, graduados y estudiantes de cursos de investigación en teoría de la probabilidad, control óptimo y matemática aplicada, buscando la mejor comprensión de la teoría; expertos en procesos de decisión markovianos, investigadores profesionales y amateurs. Los investigadores activos también pueden referirse a este libro como la aplicación de métodos matemáticos y teoremas.