DECSAI
Departamento de Ciencias de la Computación e I.A.
Universidad de Granada

Aplicaciones de la programación dinámica

Fernando Berzal, berzal@acm.org

9 de mayo de 2024

Resumen

La **programación dinámica** es un método matemático de optimización y paradigma algorítmico desarrollado por Richard Ernest Bellman en los años 50 del siglo XX. Se puede utilizar cuando la solución de un problema se puede interpretar como el resultado de una sucesión de decisiones.

Como método matemático de optimización...

En problemas matemáticos de optimización, con programación dinámica, se simplifica un problema descomponiéndolo en una secuencia de decisiones:

- Se define una secuencia de funciones $V_1, V_2, ..., V_n$ sobre los estados del problema: $V_i(y)$ es el valor obtenido para el estado y en el paso i.
- $V_n(y)$ es el valor obtenido para el estado y en el paso final n, que nos dará la solución óptima del problema.
- Los valores V_i en pasos anteriores $i = 1, 2, ..., n-2, n-1, n$ se encuentran utilizando una expresión recursiva denominada ecuación de Bellman.
- Cuando V_{i-1} ya se ha calculado para los estados de interés, se puede obtener V_i de forma directa.
- Los valores óptimos para las variables de decisión pueden recuperarse, uno a uno, observando los cálculos realizados para obtener el valor óptimo.

División optimal

Divida una cantidad positiva x en n partes de forma que el producto de esas n partes sea máximo.

Llamemos $f_n(x)$ al máximo producto alcanzable. Si c es el valor de la primera subdivisión, $x-c$ es el resto (la suma de los valores de las $n-1$ partes restantes).
La ecuación recursiva que describe este problema es:

$$f_n(x) = \max_{0\leq c\leq x}\{cf_{n-1}(x-c)\}$$

A partir de la ecuación recursiva, calculamos iterativamente los valores de la solución óptima:

- Para $n=1$,
$$f_1(x) = x$$
Por tanto, $f_1(x-c) = x-c$.

- Para $n=2$,
$$f_2(x) = \max_{0\leq c\leq x}\{cf_1(x-c)\} = \max_{0\leq c\leq x}\{c(x-c)\}$$
El máximo se obtiene para $c = x/2$. La política óptima, en consecuencia, es $\{x/2, x/2\}$ y el valor máximo del producto es $f_2(x) = (x/2)^2$.

- Para $n=3$,
$$f_3(x) = \max_{0\leq c\leq x}\{cf_2(x-c)\} = \max_{0\leq c\leq x}\{c(x-c)^2/4\}$$
El máximo se obtiene para $c = x/3$. La política óptima, en consecuencia, es $\{x/3, x/3, x/3\}$ y el valor máximo del producto es $f_3(x) = (x/3)^3$.

- Dados los resultados anteriores, conjeturamos que la política óptima para un n arbitrario es $\{x/n, x/n, .., x/n\}$ y el valor máximo del producto es $f_n(x) = (x/n)^n$, lo que podemos demostrar por inducción.

Demostración. Por inducción:

Para $n=k$:
$$f_k(x) = (x/k)^k$$

Para $n=k+1$:
$$f_{k+1}(x) = \max_{0\leq c\leq x}\{cf_k(x-c)\} = \max_{0\leq c\leq x}\{c(x-c)^k/k^k\}$$

$$\frac{df_{k+1}}{dc} = \frac{(x-c-kc)(x-c)^{k-1}}{k^k} = 0$$

$$x-c-kc = 0$$

$$x = (k+1)c$$

$$c = \frac{x}{k+1}$$

$$f_{k+1}(x) = \frac{x}{k+1}\left(\frac{kx}{k+1}\right)^k \frac{1}{k^k} = \left(\frac{x}{k+1}\right)^{k+1}$$

□

En teoría de control...

Diseño de sistemas de control	Formulación de la Mecánica Clásica
Cálculo de variaciones	Ecuación de Euler-Lagrange
Principio del máximo de Pontryagin	Principio de Hamilton
Programación dinámica	Teoría de Hamilton-Jacobi

Típicamente, en teoría de control, queremos encontrar una señal de control admisible u^* que consiga que un sistema $dx(t)/dt = g(x(t), u(t), t)$ siga una trayectoria admisible x^* en un intervalo de tiempo continuo que minimice una función de coste

$$J = b(x(t_1), t_1) + \int_{t_0}^{t_1} f(x(t), u(t), t)dt$$

La solución a este problema es una política de control óptima u^* que produce una trayectoria óptima x^* con un coste óptimo J^* que obecede la ecuación de Hamilton-Jacobi-Bellman (a.k.a. ecuación fundamental de la programación dinámica):

$$-J_t^* = \min_u \left\{ f(x(t), u(t), t) + {J_x^*}^\top g(x(t), u(t), t) \right\}$$

donde

$$J_t^* = \frac{\partial J^*}{\partial t}$$

y

$$J_x^* = \frac{\partial J^*}{\partial x} = \left[\frac{\partial J^*}{\partial x_1} \; \frac{\partial J^*}{\partial x_2} \; ... \; \frac{\partial J^*}{\partial x_n} \right]^\top$$

Una forma de resolver el problema es minimizar u en términos de t, x y la función J_x^*. El resultado se incluye en la ecuación de Hamilton-Jacobi-Bellman y la ecuación diferencial parcial se resuelve numéricamente con la condición límite $J(t_1) = b(x(t_1), t_1)$.

Alternativamente, el proceso continuo se aproxima con un sistema discreto, lo que nos proporciona una recurrencia análoga a la ecuación de Hamilton-Jacobi-Bellman, conocida como ecuación de Bellman:

$$J_k^*(x_{n-k}) = \min_{u_{n-k}} \left\{ \hat{f}(x_{n-k}, u_{n-k}) + J_{k-1}^*(\hat{g}(x_{n-k}, u_{n-k})) \right\}$$

en la etapa k de n intervalos discretos de tiempo equiespaciados, siendo $\hat{f}$ y $\hat{g}$ las aproximaciones discretas de f y g.

La ecuación de Bellman puede resolverse para encontrar la solución exacta de la aproximación discreta de nuestro problema de optimización original.

Catenarias, cicloides y curvas braquistócronas

Una curva braquistócrona (del griego brachistos, el más corto, y chronos, intervalo de tiempo), o curva del descenso más rápido, es la curva entre dos puntos que es recorrida en menor tiempo por un cuerpo que comienza en el punto inicial con velocidad cero, y que debe desplazarse a lo largo de la curva hasta llegar al segundo punto, bajo acción de una fuerza de gravedad constante y suponiendo que no existe fricción.

Curiosidad

Newton resolvió el problema de la braquistócrona, el primer resultado del cálculo de variaciones, el 24 de enero de 1697.

El principio de conservación de la energía requiere que la velocidad de un cuerpo en un campo gravitatorio uniforme venga dada por:

$$\frac{1}{2}mv^2 = mgy$$

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

donde y es la altura a la que se encuentra el cuerpo.
El espacio recorrido viene dado por

$$S = \int_{x_a}^{x_b} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

El tiempo entre dos puntos a y b viene dado por

$$\Delta t = \int_{s_a}^{s_b} \frac{ds}{v} = \int_{x_a}^{x_b} \sqrt{\frac{1 + y'(x)^2}{2gy(x)}} dx$$

Por comodidad, hacemos $x_a = 0$ y $x_b = x_n$, por lo que tenemos que minimizar la siguiente función de coste (el tiempo que se tarda en recorrer la curva):

$$J(y(x)) = \int_0^{x_n} \sqrt{\frac{1 + y'(x)^2}{2gy(x)}} dx = \int_0^{x_n} t(y(x))dx$$

Discretizando, podemos obtener una solución aproximada del problema original:

$$J_k^*(y(x_{n-k})) = \min_{y(x_{n-k})} \{t(y(x_{n-k})) + J_{k-1}^*(y(x_{n-k}))\}$$

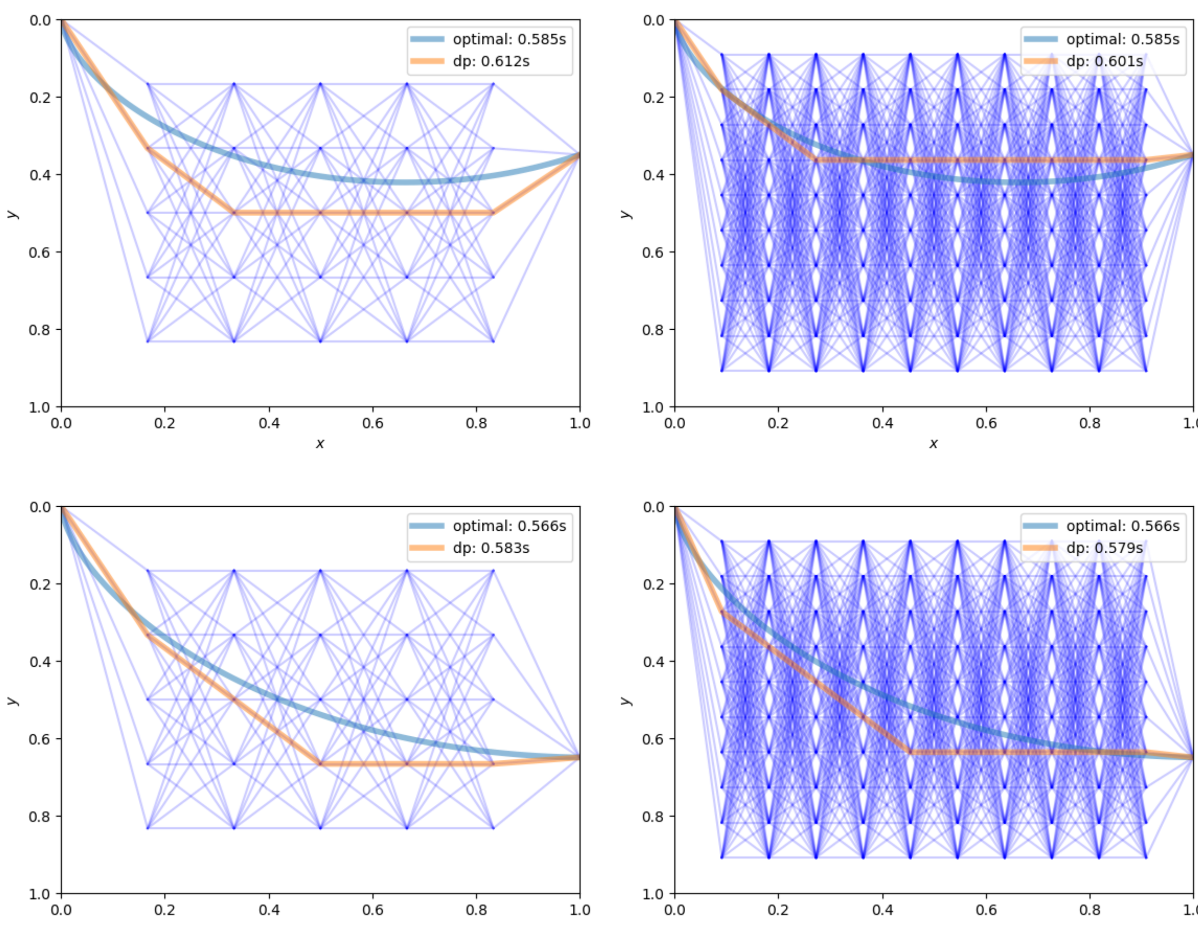
Derivación alternativa: El tiempo necesario para ir de un punto P a un punto Q lo podemos estimar de la siguiente forma:

- Aplicando el principio de conservación de la energía: $mgy_P + \frac{1}{2}mv_P^2 = mgy_Q + \frac{1}{2}mv_Q^2$

- El incremento de velocidad en Q es $v_Q - v_P = \sqrt{2g(y_P - y_Q)}$.
- La distancia de P a Q es $s_{PQ} = \sqrt{(x_P - x_Q)^2 + (y_P - y_Q)^2}$
- Cuando la aceleración es constante, la velocidad media de la partícula en el plano inclinado entre P y Q es la media de sus velocidades en los extremos: $v_{PQ} = \frac{v_P + v_Q}{2}$
- El tiempo empleado en ir de P a Q no es más que el cociente entre la longitud del plano inclinado s_{PQ} y la velocidad media v_{PQ}: $t_{PQ} = \frac{s_{PQ}}{v_{PQ}}$

  NOTA: La velocidad en cada momento sólo depende de la altura a la que nos encontremos, siendo v_P la velocidad inicial y v_Q la velocidad final el tramo PQ. El tiempo exacto que tardamos en recorrer el plano inclinado (sin rozamiento) sería entonces $t_{PQ} = \frac{2 s_{PQ}}{v_P + v_Q} = \frac{2}{\sqrt{2g}} \frac{s_{PQ}}{\sqrt{y_P} + \sqrt{y_Q}}$

Discretizando el espacio que tenemos que recorrer (de 0 a x_n en el eje x y de 0 a $-y_{max}$ en el eje y), podemos calcular el tiempo necesario para recorrer la distancia entre cada par de puntos adyacentes en el eje x y encontrar el camino más corto entre el punto inicial y el punto final de la curva. En la figura de debajo se puede ver el resultado utilizando 5 y 10 puntos intermedios en la trayectoria, junto con la curva óptima, para dos cambios de altura diferentes...



Curiosidad

Analíticamente, se puede obtener la curva cicloide que resuelve el problema: el cicloide inverso generado por un círculo de diámetro $D = 2r$. Su especificación paramétrica es:

$$\begin{cases} x = r(\theta - \sin\theta) \\ y = r(1 - \cos\theta) \end{cases}$$

donde $r = 1/4gC^2$ y $C = 1/\sqrt{2gy(1 + y'^2)}$,
En el problema del braquistócrono, el movimiento del cuerpo viene dado por la evolución del parámetro $\theta(t) = \omega t$, con velocidad angular $\omega = \sqrt{g/r}$, donde t es el tiempo transcurrido desde que el cuerpo se soltó en el punto (0,0).

Control óptimo de un coche de Fórmula 1

Deseamos encontrar la trayectoria óptima que ha de seguir un vehículo dentro de un circuito para llegar de un punto a otro en el menor tiempo posible.

La misma estrategia que utilizamos para resolver el problema del braquistócrono se puede utilizar para optimizar la trayectoria de un vehículo de carreras.
Igual que antes, intentamos minimizar el tiempo empleado en llegar de un punto a otro. Pero ahora tenemos que incorporar a nuestra función de coste las penalizaciones asociadas a violar las leyes de la física (circular fuera del circuito) o los límites del coche (cambios bruscos de velocidad o de ángulo quedan descartados). Hemos de considerar, por ejemplo, la fuerza necesaria para cambiar de un estado a otro: la asociada al movimiento (masa por aceleración, dada por la Ley de Newton) más la fuerza aerodinámica que actúa sobre el coche (limitada por la fricción de la carretera y la potencia del coche). Si el movimiento requiere una fuerza superior a la soportada por el vehículo, deja de ser factible, por lo que conlleva una penalización elevada.
En la discretización del problema, dividimos la calzada en la que se realiza la carrera en múltiples etapas (y distintos puntos en los que nos podemos encontrar en cada etapa, dentro del ancho de la calzada). En cada posición, sabemos cuáles son sus coordenadas, pero no la velocidad con la que puede llegar el vehículo a esa posición, para lo que necesitamos calcular trayectorias completas.
Para obtener la trayectoria óptima, calculamos el camino mínimo entre el punto inicial y el final teniendo en cuenta los costes asociados a cada transición (penalizaciones incluidas)...
Al final, la trayectoria óptima calculada hará que el vehículo se abra antes de cada curva para poder tomarla con la máxima velocidad posible, acelerando al máximo en las rectas y reduciendo la velocidad antes de las curvas (más cuanto más cerradas sean éstas).

Curiosidad

Exactamente la misma estrategia basada en programación dinámica se puede usar también para planificar la trayectoria de un vehículo autónomo en una ciudad. Por ejemplo, en calles de doble sentido, girar a la izquierda puede suponer una maniobra arriesgada. Las penalizaciones asociadas a dicha maniobra pueden conducir a que el coche, en lugar de girar directamente a la izquierda en busca de su destino, prefiera seguir el línea recta y rodear una manzana a la derecha para encarar su destino en línea recta, sin necesidad de realizar giros a la izquierda.

Aplicaciones en Economía

Ahorro óptimo

Deseamos maximizar nuestro bienestar futuro, para lo cual caracterizamos nuestro nivel de consumo actual con una función de utilidad. Tenemos que elegir, en cada momento, qué cantidad de nuestros recursos consumir y cuánto ahorrar de cara al futuro (invirtiendo ese ahorro en capital). Es lo que los economistas denominan una elección intertemporal...
Partimos de un capital inicial k_0 y el consumo futuro se descuenta con una tasa fija $\beta \in (0, 1)$ (esto es, el consumo actual es preferible al consumo futuro). Asumiendo que viviremos N años, ¿cuál es la estrategia de ahorro que maximiza nuestro bienestar futuro?

- Frank P. Ramsey: A Mathematical Theory of Saving, The Economic Journal, 1928. URL https://www.jstor.org/stable/2224098

Denominamos c_t a nuestro consumo en el período t.
Asumimos que la utilidad asociada a nuestro consumo es $u(c_t) = \ln(c_t)$, que satisface la ley de la utilidad marginal decreciente (cuanto mayor es la cantidad que consumimos de un bien, menor es la utilidad marginal que nos aporta cada nueva unidad del mismo). Además, se utiliza un descuento para la utilidad futura: se descuenta la utilidad futura con un factor b, $0 < b < 1$, por cada período de tiempo.
El capital acumulado se aproxima por $k_{t+1} = f(k_t) - c_t$, donde f es una función de producción que satisface las condiciones de Inada (podemos interpretarla como la rentabilidad que obtenemos al invertir nuestro capital). Para simplificar, asumamos que $f(k_t) = Ak_t^a$, con A constante positiva y $0 < a < 1$. Entonces, $k_{t+1} = Ak_t^a - c_t$.
Con las suposiciones anteriores, el problema del consumidor se plantea como

$$\text{máx} \sum_{t=0}^{N} b^t \ln(c_t)$$

sujeto a la restricción

$$k_{t+1} = Ak_t^a - c_t \geq 0$$

El capital inicial, k_0, nos viene dado, pero tenemos que resolver el problema para todos los consumos c_t, con $0 \leq t \leq N$.
Mediante programación dinámica, descomponemos nuestro problema en una secuencia de decisiones. Para ello, definimos una secuencia de funciones $V_t(k)$ que representan el valor de tener una cantidad de capital k en el período de tiempo t. Asumimos también que no tiene utilidad tener capital después de nuestra muerte, por lo que $V_{N+1}(k) = 0$.
El valor de cualquier cantidad de capital en cualquier instante de tiempo puede calcularse hacia atrás usando la siguiente ecuación de Bellman:

$$V_t(k_t) = \max\{\ln(c_t) + bV_{t+1}(k_{t+1})\}$$

Este problema es mucho más sencillo de resolver que el original, ya que sólo intervienen dos variables, k_{t+1} y c_t, que además están relacionadas: $k_{t+1} = Ak_t^a - c_t \geq 0$.
$V_{N+1}(k)$ ya lo conocemos (0), por lo que mediante la ecuación de Bellman podemos calcular $V_N(k)$. Esto nos permite calcular $V_{N-1}(k)$ y así sucesivamente hasta llegar a $V_0(k)$.
Podemos obtener $V_{N-j}(k)$ como el máximo de $\ln(c_{N-j}) + bV_{N-j+1}(Ak^a - c_{N-j})$:

$$V_{N-j}(k) = a \sum_{i=0}^{j} a^i b^i \ln k + v_{N-j}$$

donde v_{N-j} es una constante.
La cantidad óptima de consumo es

$$c_{N-j}(k) = \frac{1}{\sum_{i=0}^{j} a^i b^i} Ak^a$$

Esto es, de acuerdo con nuestras suposiciones iniciales, la estrategia óptima es consumir una fracción cada vez más grande de nuestra riqueza conforme envejecemos, hasta consumir la riqueza restante el año N ($c_N = Ak^a$), nuestro último año de vida.

Curiosidad

Existen modelos económicos de crecimiento que se inspiran en el modelo de Ramsey. El modelo de Ramsey-Cass-Koopsmans, basado en el trabajo de Frank Ramsey y extendido por David Cass y Tjalling Koopsmans, intenta explicar el crecimiento económico a largo plazo, al margen de perturbaciones como ineficiencias del mercado, shocks exógenos o desigualdades entre hogares...
Para maximizar la función de utilidad asociada al consumo agregado de la sociedad, y que el problema sea tratable analíticamente, asume que la sociedad está compuesta por “individuos inmortales idénticos con funciones de utilidad no cambiantes” [sic].

Bibliografía

- Nancy L. Stokey , Robert E. Lucas Jr. , Edward C. Prescott: Recursive Methods in Economic Dynamics, Harvard University Press, 1989, ISBN 0674750969.

- Lars Ljungqvist, Thomas J. Sargent: Recursive Macroeconomic Theory, 4th edition, MIT Press, 2018. ISBN 0262038668.
- Jianjun Miao: Economic Dynamics in Discrete Time, 2nd edition, MIT Press, 2020. ISBN 0262043629.
- Avinash K. Dixit, Robert S. Pindyck: Investment under Uncertainty, Princeton University Press, 1994. ISBN 0691034109
- Mario J. Miranda, Paul L. Fackler: Applied Computational Economics and Finance, MIT Press, 2002. ISBN 0262134209.
- Patrick L. Anderson: Business Economics and Finance with Matlab, GIS, and Simulation Models, Chapman and Hall / CRC Press, 20043. ISBN 1584883480.