Departamento de Estadística y Matemática

Documento de Trabajo Nº 6

Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

Programación Dinámica.
La Ecuación de Bellman y el Teorema de la Envolvente

Jorge Mauricio Oviedo [1]

Resumen: El presente trabajo propone resaltar la importancia de las técnicas de Optimización Dinámicas y dentro de ellas, la Programación Dinámica y el teorema de la Envolvente. Dado el amplio y creciente uso de dichas técnicas en la Dinámica Económica, como así también la dificultad de acceder a dichos tópicos, se desarrollan las ideas de una manera clara e intuitiva y sin perdida de rigor matemático. Se intenta con ello llenar el vació existente en cuanto a bibliografía accesible al estudiante de grado.

Palabras clave: Optimización Dinámica, Programación Dinámica, Función de Valor, Ecuación de Bellman, Teorema de la Envolvente.

[1] joviedo@eco.unc.edu.ar

1.- Introducción

La optimización Dinámica, como su nombre lo indica, estudia la optimización de sistemas dinámicos, es decir, sistemas que evolucionan en el tiempo. De esta manera, se trata de guiar o controlar el sistema de manera óptima a lo largo de un horizonte temporal dado, de acuerdo a un objetivo previamente fijado.

En la literatura económica actual se habla cada vez más y con más insistencia de Economía Dinámica, siendo esta técnica de difícil abordaje por parte del estudiante de grado. Dentro de la Optimización Dinámica se destaca el uso en economía de las técnicas de la programación dinámica y la ecuación de Bellman.

Dado el amplio y creciente uso de dichas técnicas en la Dinámica Económica, como así también la dificultad de acceder a dichos tópicos, el propósito de este ensayo es desarrollar las ideas de una manera clara e intuitiva y sin pérdida de rigor matemático. Se intenta con ello llenar el vació existente en cuanto a bibliografía accesible al estudiante de pregrado.

El trabajo tiene la siguiente estructura: en la sección dos se describen las principales diferencias entre la optimización dinámica y la estática para posteriormente en la sección tres, pasar a describir el problema en términos matemáticos. La sección cuatro se brindan las definiciones y teoremas claves para en la siguiente abordar una solución a estos problemas por medio del Teorema de la Envolvente.

2.- Optimización Estática versus Dinámica

En los problemas de optimización estática el objetivo consistía en determinar el o los valores de una o varias variables que hacían máxima una función objetivo en un determinado momento del tiempo. En la optimización dinámica, en cambio, se trata de resolver una sucesión (finita o infinita) o un continuo de problemas de optimización estática pero con la gran diferencia de que estos problemas no pueden resolverse por separado y de manera independiente por las siguientes dos razones:

1.- La función objetivo no depende de valores independientes de las variables en cada momento del tiempo sino que generalmente están interrelacionadas en la forma que aportan valor al funcional objetivo de tal manera que lo óptimo se define en términos de óptimo para todo el horizonte de análisis y no en cada instante en particular y por separado.

2.- En segundo lugar, la sucesión de problemas y de las variables de elección (que en estos problemas se denominan variables de control) se hallan interconectados en el tiempo por lo que se conocen como "***ecuaciones de movimiento***" o evolución de las variables de estado. Estas variables describen en cada instante de tiempo la evolución del sistema y dicha evolución está influida por las variables de control.

De esta manera, las variables de control en este caso serían análogas a las variables de decisión en el problema de optimización estática y a diferencia de tal problema, estas variables no sólo influyen en la función objetivo sino que también influyen en la evolución dinámica del sistema vía las ecuaciones de movimiento de las variables de estado.

En este sentido, el objetivo consiste en "controlar" la evolución dinámica del sistema o, lo que es lo mismo, en determinar un continuo o una secuencia (finita o infinita) de variables de modo tal que hagan máximo un funcional objetivo teniendo en cuenta esta doble influencia de los controles.

En cuanto a su tratamiento matemático, los problemas de Optimización Dinámica pueden resolverse por medio de algunas de estas las siguientes técnicas:

1.- El Cálculo de Variaciones

2.- El Control Óptimo

3.- La Programación Dinámica

Si bien las tres técnicas permiten abordar problemas en tiempo discreto y continuo, los métodos 1 y 2 se utilizan generalmente para el tiempo continúo y el restante para el tiempo discreto.

A continuación se desarrollará la técnica de Programación Dinámica por ser la de mayor uso en Economía.

3.- Descripción Matemática del Problema – Definición de variables

Considérese el siguiente problema general:

$$\max_{\{x_t\}} \sum_{t=0}^{\infty} \beta^t f(x_t, w_t)$$
$$st:\ w_{t+1} = \Gamma(x_t, w_t)$$

en que:

x_t es la variable de control por cuyo intermedio se debe lograr maximizar el funcional objetivo. Este último viene descrito por la sumatoria infinita.

w_t es la variable de estado. La misma describe la evolución del sistema dinámico vía la ecuación de movimiento.

β es el factor de descuento que pondera los aportes al funcional vía f en los distintos momentos del tiempo (β<1).

Obsérvese cómo funcionan las interrelaciones entre las variables: supongamos que partimos del periodo t = 0 en donde el valor de w esta fijado en w_o. Una vez elegido el control de una manera optima para t=0 se determinará el aporte al funcional vía f por un valor de $f(w_o,x_o)$ y se determinará el valor del estado en t=1.

Nótese como la elección de x no solo condicionó el valor de f si no que también afectó el valor del estado en el periodo siguiente repercutiendo éste en el funcional pues, ya comienza a aportar a f en el periodo t =1. De esta manera, se entiende como la elección de los controles en cada período afecta las elecciones y las valoraciones en los periodos sucesivos siguientes.

Otra cuestión a tener en cuenta es la interpretación de las variables de estado. Éstas de alguna manera determinan el nivel del sistema en cada instante, son una especie de indicador que establecen en cada etapa el valor de arranque, el nivel de recursos con que cuenta el sistema en el resto de su trayectoria hasta el final. Así wt significa que el sistema cuenta con wt para el resto de la trayectoria desde el periodo actual hasta el final. La elección que se haga de x en t determinará w_{t+1} en t+1 siendo esto una medida de cuánto le resta al sistema desde t+1 hasta el final vía la ecuación de movimiento.

4.- La Función de Valor y la Ecuación de Bellman

La función de Valor o función indirecta, $J(w_t)$, se define como el valor máximo que puede alcanzar la función objetivo, una vez que se han seleccionado de manera óptima los controles desde el momento t hasta el final del problema y partiendo desde el estado inicial w_t.

$$J(w_t) = \max_{\{x_h\}} \sum_{h=t}^{\infty} \beta^t f(x_h, w_h)$$
$$sa: \; w_{h+1} = \Gamma(x_h, w_h) \qquad h = t, t+1, ...$$

Intuitivamente, se puede pensar la funcionalidad de J con respecto a w_t como una función de las condiciones iniciales (y por ende temporales, dada la característica dinámica del sistema), es decir la función de Valor otorga el valor máximo del funcional para distintos subproblemas que parten de condiciones “iniciales-temporales variables” hasta el final.

La Ecuación de Bellman: esta ecuación es una relación recursiva fundamental que traduce matemáticamente el principio Básico de la Programación Dinámica llamado el principio de Optimalidad de Bellman (1957) y que se enuncia como sigue:

"Un política óptima tiene la propiedad de que, cualesquiera sean el estado y las decisiones iniciales tomadas (es decir, el control), las restantes decisiones deben constituir una política óptima con independencia del estado resultante de la primer decisión"[2].

En términos Matemáticos el Principio de la Optimalidad se puede expresar por medio de lo que se conoce como la relación de recurrencia fundamental de la programación dinámica o **Ecuación de Bellman** así:

$$J(w_t) = \max_{x_t} \left[f(x_t, w_t) + \beta J(w_{t+1}) \right]$$

$$sa: \; w_{t+1} = \Gamma(x_t, w_t)$$

la que coloquialmente dice que el valor máximo que se puede obtener desde el estado w_t es el valor máximo desde el estado siguiente más el valor máximo de f una vez optimizada con respecto a la variable de control para el periodo t.

Esta ecuación en apariencia obvia, dada la forma de definir la función de valor, posee un valor incalculable en el sentido de que permite transformar un problema de múltiples periodos en múltiples problemas de un solo periodo[3]. Sin embargo cuando se trata de problemas con horizonte temporal de optimización infinito su poder se encuentra en que habilita la aplicación de métodos recursivos para hallar las soluciones explicitas[4]. Independientemente del horizonte de optimización, la ecuación de Bellman puede utilizarse con otras herramientas matemáticas para hallar descripciones cualitativas y numéricas explicitas[5] que son ampliamente usadas en los problemas de optimización intertemporal en Economía. A continuación se expone este último método.

Siguiendo con la notación anterior e introduciendo la ecuación de movimiento de la variable de estado en la ecuación de Bellman se tiene:

---

[2] La demostración de la necesidad del principio de optimalidad se obtiene inmediatamente por contradicción. Aris (1964), expresa el principio así: *"Si tú no haces lo mejor con lo que llegues a tener, nunca harás lo mejor que tú podrías haber hecho con lo deberías haber tenido"*.

[3] Esta ecuación es suficiente para resolver problemas de etapas finitas. Véase Kreeps (1995) Apéndice 2 para un excelente ejemplo de la aplicación de esta técnica.

[4] Como los métodos que iteran la función de Valor o los que iteran la función de política óptima.

$$J(w_t) = \max_{x_t} \left\{ f(x_t, w_t) + \beta J\left[\Gamma(x_t, w_t)\right] \right\}$$

De la Condición de Primer Orden surge:

$$f_{x_t}(x_t, w_t) + \beta J'\left[\Gamma(x_t, w_t)\right]\Gamma_{x_t}(x_t, w_t) = 0$$

Observando con detenimiento la expresión anterior se puede notar que el término J'[.] es desconocido funcionalmente por lo que necesitamos valernos de algún teorema que haga el "milagro" de cambiar J' por algo conocido. Por suerte Dios existe y existe también este teorema capaz de hacer esta brillante pirueta.

5.- <u>Teorema de la Envolvente</u>

Recordemos que lo que se desea conocer es la expresión funcional de

$$J'\left[\Gamma(x_t, w_t A_t)\right] = J_\Gamma\left[\Gamma(x_t, w_t A_t)\right] = \frac{\partial J(\Gamma)}{\partial \Gamma} = \frac{\partial J(w_{t+1})}{\partial w_{t+1}}$$

Para ello volvamos a la condición de primer orden de cuya solución resultará el valor óptimo de la variable de control como función del estado del sistema en ese momento. Así se tiene que:

$$x_t^* = x_t(w_t)$$

Sustituyendo el óptimo en la ecuación de Bellman resulta:

$$J(w_t) = f[x_t(w_t), w_t] + \beta J\{\Gamma[x_t(w_t), w_t]\}$$

Derivando la expresión anterior con respecto a w_t se arriba a:

[5] Aunque éstas últimas con mayor dificultad en algunos casos.

$$\frac{\partial J(w_t)}{\partial w_t} = f_{x_t}[x_t(w_t), w_t]x'_t(w_t) + f_{w_t}(\cdot) + \beta J_\Gamma(\cdot)\{\Gamma_{x_t}(\cdot)x'_t(w_t) + \Gamma_{w_t}(\cdot)]\}$$

despejando f_{xt} desde la condición de Primer Orden y remplazando en la expresión anterior resulta:

$$\frac{\partial J(w_t)}{\partial w_t} = f_{w_t}(x_t, w_t) + \beta J_\Gamma[(x_t, w_t)]\Gamma_{w_t}(x_t, w_t)$$

Este resultado era previsto por lo que se conoce como **Teorema de la Envolvente** el cual establece que la derivada de la función de Valor con respecto a uno de los parámetros del problema en el óptimo es simplemente la derivada de la función directa con respecto al parámetro evaluada en el óptimo, siendo despreciable los efectos indirectos sobre la función objetivo.

Volviendo a la CPO y despejando de ésta se tiene

$$\beta\frac{\partial J(w_t)}{\partial w_t} = -\frac{f'(x_t)}{\Gamma_{x_t}(x_t, w_t)}$$

Reemplazando en lo obtenido por el Teorema de la Envolvente se arriba a:

$$\frac{\partial J(w_t)}{\partial w_t} = f_{w_t}(x_t, w_t) - \frac{f'(x_t + w_t)\Gamma_{w_t}(x_t, w_t)}{\Gamma_{x_t}(x_t, w_t)}$$

Recordando que lo deseado era: $\frac{\partial J(w_{t+1})}{\partial w_{t+1}}$ se deduce que[6]

[6] Reemplácese t por t + 1 en los índices de las variables en las cuales se evalúan las respectivas funciones. Téngase en cuenta además que este procedimiento es válido ya que se pueden replantear simultáneamente todos estos desarrollos en la siguiente formulación de la ecuación de Bellman, siendo ésta igualmente legítima:

$$\frac{\partial J(w_{t+1})}{\partial w_{t+1}} = f_{w_t}(x_{t+1}, w_{t+1}) - \frac{f'(x_{t+1}, w_{t+1})\Gamma_{w_t}(x_{t+1}, w_{t+1})}{\Gamma_{x_t}(x_{t+1}, w_{t+1})}$$

Reemplazando la expresión anterior en la condición de primer orden se obtiene la siguiente expresión final:

$$f'(x_t, w_t) - \beta[f_{w_t}(x_{t+1}, w_{t+1}) - \frac{f'(x_{t+1}, w_{t+1})\Gamma_{w_t}(x_{t+1}, w_{t+1})}{\Gamma_{x_t}(x_{t+1}, w_{t+1})}]\Gamma_{x_t}(x_t, w_t) = 0$$

De esta manera vía la aplicación del Teorema de la Envolvente se consiguió sustituir el término $J'[.]$ por una expresión conocida en términos de los parámetros funcionales del problema original.

Esta relación es conocida como la ecuación de Ëuler[7] para la variable x_t. Esta ecuación provee de gran riqueza cualitativa en el análisis de los modelos dinámicos en economía pues describe el comportamiento de las sendas óptimas en términos de relaciones de sustitución entre un par de períodos consecutivos de tiempo.

Por otra parte se pueden obtener las trayectorias óptimas de las variables de elección (x_t y w_t) que maximizan intertemporalmente la función objetivo, es decir se pueden hallar x(t) y w(t) para observar su evolución en cualquier momento del tiempo. Para ello se debe hacer uso de la ecuación de transición a la hora de arribar al siguiente Sistema de Ecuaciones en Diferencias de Primer Orden

$$\begin{cases} f'(x_t, w_t) - \beta[f_{w_t}(x_{t+1}, w_{t+1}) - \dfrac{f'(x_{t+1}, w_{t+1})\Gamma_{w_t}(x_{t+1}, w_{t+1})}{\Gamma_{x_t}(x_{t+1}, w_{t+1})}]\Gamma_{x_t}(x_t, w_t) = 0 \\ w_{t+1} = \Gamma(w_t, x_t) \end{cases}$$

$$J(w_{t+1}) = \max_{x_{t+1}}\left[f(x_{t+1}, w_{t+1}) + \beta J(w_{t+2})\right]$$

$$sa: \; w_{t+2} = \Gamma(x_{t+1}, w_{t+1})$$

[7] Si bien las Ecuaciones de Euler, que es la condición de Primer Orden cuando el problema se analiza desde el Enfoque del Cálculo de Variaciones y que datan desde el Siglo XVIII, se describieron originalmente para el caso de tiempo continuo, en Economía se readaptan las mismas al caso discreto siendo las expuestas en este escrito una de las formas de hacerlo.

Cuya resolución permite hallar las trayectorias deseadas[8].

[8] Se requerirán además las condiciones iniciales y de transversalidad provistas en todo problema de optimización dinámica.

BIBLIOGRAFÍA

Aris, R. (1961) "*The Optimal Design of Chemical Reactors*". New York, Academic Press Icc.

Bellman, Richard (1957): *"Dynamic Programing"* Princeton University Press. Princeton, New Jersey.

Cerdá Tena, Emilio (2001): *Optimización Dinamica.* Prentice Hall. España.

Intriligator, Michael D (1971)**.** "*Mathematical optimization and economic theory*". Prentice-Hall.

Kreps 1995, "*Curso de Teoría Microeconómica*". McGraw-Hill. Madrid. España

Obstfeld, Maurice and Rogoff, kenneth: "*Fundations of International Macroeconomics***".** The MIT Press. Cambridge, Massachusetts. London, Ingland.

Sargent, Thomas J. (2000): "*Recursive Macroeconomic Theory".* Second edition. Massachusetts Institute of Technology.

Stokey, Nancy and Lucas, Robert (1987): *"Recursive Methods in Economic Dynamic"* Harvard University Press.