Integrated Intelligent Research (IIR)
International Journal of Computing Algorithm
Volume: 03 Issue: 03 December 2014 Pages: 277-279
ISSN: 2278-2397

Bellman Equation in Dynamic Programming

K.Vinotha
Associate Professor, Sri ManakulaVinayagar College of Engineering, Madagadipet, Puducherry
E-mail :kvinsdev@gmail.com

Abstract - The unifying purpose of this paper to introduces basic ideas and methods of dynamic programming. It sets out the basic elements of a recursive optimization problem, describes Bellman's Principle of Optimality, the Bellman equation, and presents three methods for solving the Bellman equation with example.

Key words:Dynamic Programming, Bellmen's equation, Contraction Mapping Theorem, Blackwell's Sufficiency Conditions.

I. PRELIMINARIES - DYNAMIC PROGRAMMING

Dynamic programming is a mathematical optimization method. It refers to simplifying a complicated problem by breaking it down into simpler subproblems in a recursive manner. While some decision problems cannot be taken apart this way, decisions that span several points in time do often break apart recursively; Bellman called this the "Principle of Optimality". If subproblems can be nested recursively inside larger problems, so that dynamic programming methods are applicable, then there is a relation between the value of the larger problem and the values of the subproblems. In the optimization literature this relationship is called the Bellman equation.

II. THE BELLMAN EQUATION AND RELATED THEOREMS

A derivation of the Bellman equation

Let us consider the case of an agent that has to decide on the path of a set of control variables, $\{y_t\}_{t=0}^{\infty}$ in order to maximize the discounted sum of its future payoffs, $u(y_t, x_t)$ where x_t is state variables assumed to evolve accordingto $x_{t+1} = h(x_t, y_t), x_0$ given.We may assume that the model is markovian, a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker. The optimal value our agent can derive from this maximization process isgiven by the value function

$$V(x_t) = \max_{\{y_{t+s} \in \mathscr{D}(x_{t+s})\}_{s=0}^{\infty}} \sum_{s=0}^{\infty} \beta^s u(y_{t+s}, x_{t+s}) \qquad (1)$$

where D is the set of all feasible decisions for the variables of choice. Notethat the value function is a function of the state variable only, as since themodel is markovian, only the past is necessary to take decisions, such that allthe path can be predicted once the state variable is observed. Therefore, thevalue in *t*is only a function of x_t(1) may now be rewritten as

$$V(x_t) = \max_{\{y_t \in \mathscr{D}(x_t), \{y_{t+s} \in \mathscr{D}(x_{t+s})\}_{s=1}^{\infty}} u(y_t, x_t) + \sum_{s=1}^{\infty} \beta^s u(y_{t+s}, x_{t+s}) \qquad (2)$$

making the change of variable *k = s - 1*, (2) rewrites

$$V(x_t) = \max_{\{y_t \in \mathscr{D}(x_t), \{y_{t+1+k} \in \mathscr{D}(x_{t+1+k})\}_{s=0}^{\infty}} u(y_t, x_t) + \sum_{k=0}^{\infty} \beta^{k+1} u(y_{t+1+k}, x_{t+1+k})$$

or

$$V(x_t) = \max_{\{y_t \in \mathscr{D}(x_t)} u(y_t, x_t) + \beta \max_{\{y_{t+1+k} \in \mathscr{D}(x_{t+1+k})\}_{k=0}^{\infty}} \sum_{k=0}^{\infty} \beta^k u(y_{t+1+k}, x_{t+1+k})$$

Note that, by definition, we have

$$V(x_{t+1}) = \max_{\{y_{t+1+k} \in \mathscr{D}(x_{t+1+k})\}_{k=0}^{\infty}} \sum_{k=0}^{\infty} \beta^k u(y_{t+1+k}, x_{t+1+k}) \cdots \qquad (3)$$

Such that (3) rewrites as

$$V(x_t) = \max_{y_t \in \mathscr{D}(x_t)} u(y_t, x_t) + \beta V(x_{t+1}) \qquad \ldots(4)$$

This is the socalled*Bellman equation* that lies at the core of the dynamicprogramming theory. With this equation are associated, in each and everyperiod*t*, a set of *optimal policy functions* for *y* and *x*, which are defined by

$$\{y_t, x_{t+1}\} \in \underset{y \in \mathscr{D}(x)}{Arg\max}\ u(y, x) + \beta V(x_{t+1}) \qquad \ldots(5)$$

Our problem is now to solve (4) for the function $V(x_t)$. This problem isparticularly complicated as we are not solving for just a point that wouldsatisfy the equation, but we are interested in finding a function that satisfiesthe equation. A simple procedure to find a solution would be the following

1. Make an initial guess on the form of the value function $V_0(x_t)$
2. Update the guess using the Bellman equation such that

$$V_{i+1}(x_t) = \max_{y_t \in \mathscr{D}(x_t)} u(y_t, x_t) + \beta V_i(h(y_t, x_t))$$

3. If $V_{i+1}(x_t) = V_i(x_t)$, then a fixed point has been found and the problem is solved, if not we go back to 2, and iterate on the process untilconvergence.

In other words, solving the Bellman equation just amounts to find the fixedpoint of the bellman equation, or introduction an operator notation, findingthefixed point of the operator *T*, such that

$$V_{i+1} = TV_i$$

where*T* stands for the list of operations involved in the computation of theBellman equation. The problem is then that of the existence and the uniqueness of this fixed-point.

Volume: 03 Issue: 03 December 2014 Pages: 277-279
ISSN: 2278-2397

Luckily, mathematicians have provided conditions forthe existence and uniqueness of a solution.

Existence and uniqueness of a solution

Definition 1 A metric space is a set *S*, together with a metric $\rho : S \times S \to \mathbb{R}^{+}$, such that for all $x, y, z \in S$:

1. $\rho(x, y) \geq 0$, with $\rho(x, y) = 0$ if and only if $x = y$,
2. $\rho(x, y) = \rho(y, x)$
3. $\rho(x, y) \leq \rho(x, y) + \rho(y, z)$

Definition 2 A sequence $\{x_n\}_{n=0}^{\infty}$ in *S* converges to $x \in S$, if for each $\varepsilon > 0$ there exists an integer N_ε such that

$$\rho(x_n, x) < \varepsilon \ for\ all \quad n \geq N_\varepsilon$$

Definition 3 A sequence $\{x_n\}_{n=0}^{\infty}$ in *S* is a Cauchy sequence if for each $\varepsilon > 0$ there exists an integer N_ε such that

$$\rho(x_n, x_m) < \varepsilon \ for\ all \quad n, m \geq N_\varepsilon$$

Definition 4 A metric space (S, ρ) is complete if every Cauchy sequence in*S* converges to a point in *S*.

Definition 5 Let (S, ρ) be a metric space and $T : S \to S$ be function mapping*S* into itself. *T* is a contraction mapping (with modulus β) if for $\beta \in (0,1)$

$\rho(T_x, T_y) \leq \beta\ \rho(x, y)$, *for all* $x, y \in S$.

We then have the following remarkable theorem that establishes the existence and uniqueness of the fixed point of a contraction mapping.

Theorem 1 (Contraction Mapping Theorem) If (S, ρ) is a complete metric space and $T : S \to S$ is a contraction mapping with modulus $\beta \in (0,1)$, then

1. *T* has exactly one fixed point $V \in S$ such that $V = T\,V$,
2. for any $V \in S$, $\rho(T^n V_0, V) < \beta^n\ \rho(V^0, V)$, with *n* = 0, 1, 2 …

Since we are endowed with all the tools we need to prove the theorem, we shall do it.

Proof: In order to prove 1., we shall first prove that if we select any sequence $\{V_n\}_{n=0}^{\infty}$, such that for each *n*, $V_n \in S$ and $V_{n+1} = T\ V_n$this sequence converges and that it converges to $V \in S$. In order to show convergence of $\{V_n\}_{n=0}^{\infty}$, we shall prove that $\{V_n\}_{n=0}^{\infty}$ is a Cauchy sequence. First of all, note that the contraction property of *T* implies that

$\rho(V_2, V_1) = \rho(TV_1, TV_0) \leq \beta\ \rho(V_1, V_0)$,

and therefore

$\rho(V_{n+1}, V_n) = \rho(TV_n, TV_{n-1}) \leq \beta\ \rho(V_n, V_{n-1}) \leq \ldots \leq \beta^n\ \rho(V_1, V_0)$ Now consider two terms of the sequence, $V_m$andV_n, $m > n$. The triangle inequality implies that

$\rho(V_m, V_n) \leq \rho(V_m, V_{m-1}) + \rho(V_{m-1}, V_{m-2}) + \ldots + \rho(V_{n+1}, V_n)$ therefore, making use of the previous result, we have

$\rho(V_m, V_n) \leq (\beta^{m-1} + \beta^{m-2} + \ldots + \beta^n)\ \rho(V_1, V_0) \leq \frac{\beta^n}{1-\beta}\rho(V_1, V_0)$ Since $\beta \in (0,1), \beta^n \to 0$ as $n \to \infty$, we have that for each $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}$ such that $\rho(V_m, V_n) < \varepsilon$. Hence $\{V_n\}_{n=0}^{\infty}$ is a Cauchy sequence and it therefore converges. Further, since we have assume that *S* is complete, V_n converges to $V \in S$.

We now have to show that *V* = *T V* in order to complete the proof of the first part. Note that, for each $\varepsilon > 0$, and for $V_0 \in S$ the triangular inequality implies

$\rho(V, TV) \leq \rho(V, V_n) + \rho(V_n, TV)$

But since $\{V_n\}_{n=0}^{\infty}$ is a Cauchy sequence, we have

$\rho(V, TV) \leq \rho(V, V_n) + \rho(V_n, TV) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$ for large enough *n*, therefore *V* = *T V*.

Hence, we have proven that *T* possesses a fixed point and therefore have established its existence. We now have to prove uniqueness. This can be obtained by contradiction. Suppose, there exists another function, say W ∈ S that satisfies *W* = *T W* . Then, the definition of the fixed point implies

$\rho(V, W) = \rho(TV, TW)$

but the contraction property implies

$\rho(V, W) = \rho(TV, TW) \leq \beta\ \rho(V, W)$

which, as $\beta > 0$ implies $\rho(V, W) = 0$ and so *V* = *W* . The limit is then unique.

Proving 2.is straightforward as

$\rho(T^n V_0, V) = \rho(T^n V_0, TV) \leq \beta\ \rho(T^{n-1} V_0, V)$

but we have $\rho(T^{n-1} V_0, V) = \rho(T^{n-1} V_0, TV)$ such that

$\rho(T^n V_0, V) = \rho(T^n V_0, TV) \leq \beta\ \rho(T^{n-1} V_0, V) \leq \beta^2\ \rho(T^{n-2} V_0, V) \leq \ldots \leq \beta^n\ \rho(V_0, V)$

which completes the proof

This theorem is of great importance as it establishes that any operator that possesses the contraction property will exhibit a unique fixed-point, which therefore provides some rationale to the algorithm we were designing in the previous section. It also insure that whatever the initial condition for thisalgorithm, if the value function satisfies a contraction property, simple iterations will deliver the solution. It therefore remains to provide conditions for the value function to be a contraction. These are provided by the following theorem.

Theorem 2(Blackwell's Sufficiency Conditions) Let $X \subseteq \mathbb{R}^{\ell}$ and *B(X)* be the space of bounded functions $V : X \to \mathbb{R}$ with the uniform metric. Let $T : B(X) \to B(X)$ be an operator satisfying

1. (*Monotonicity*) Let $V, W \in B(X)$, *if* $V(x) \leq W(x)$ *for all* $x \in X$, then $TV(x) \leq TW(x)$
2. (*Discounting*) There exists some constant $\beta \in (0,1)$ such that for all $V \in B(X)$ and $a \geq 0$, we have

$T(V + a) \leq TV + \beta a$

then*T* is a contraction with modulus β.

Proof: Let us consider two functions $V, W \in B(X)$ satisfying 1. and 2., and such that

$V \leq W + \rho(V, W)$

Monotonicity first implies that

$TV \leq T(W + \rho(V, W))$

and discounting

$TV \leq T(W + \beta\rho(V, W))$

since $\rho(V, W) \geq 0$ plays the same roles as a . We therefore get

$TV - T \leq \beta\rho(V, W)$

Likewise, if we now consider that $V \leq W + \rho(V, W)$, we end up with

$TW - TV \leq \beta\rho(V, W)$

Consequently, we have

$|TW - TV| \leq \beta\rho(V, W)$

so that

$(TW - TV) \leq \beta\rho(V, W)$

Which defines a contraction. This completes the proof.
This theorem is extremely useful as it gives us simple tools to check whether a problem is a contraction and therefore permits to check whether the simple algorithm we were defined is appropriate for the problem we have in hand.
As an example, let us consider the optimal growth model, for which the Bellman equation writes

$$V(k_t) - \max_{c_t \in \mathscr{C}} u(c_t) + \beta V(k_{t+1})$$

with$k_{t+1} = F(k_t) - c_t$. In order to save on notations, let us drop the time subscript and denote the next period capital stock by k' , such that the Bellman equation rewrites, plugging the law of motion of capital in the utility function

$$V(k) = \max_{k' \in \mathscr{R}} u(F(k) - k') + \beta V(k')$$

Let us now define the operator T as

$$TV(k) = \max_{k' \in \mathscr{R}} u(F(k) - k') + \beta V(k')$$

We would like to know if T is a contraction and therefore if there exists a unique function V such that

$V(k) = (TV)(k)$

In order to achieve this task, we just have to check whether T is monotonic and satisfies the discounting property.
Monotonicity: Let us consider two candidate value functions, V and W , such that $V(k) \leq W(k)$ for all $k \in \mathscr{R}$. What we want to show is that $TV(k) \leq (TW)(k)$. In order to do that, let us denote by $\tilde{k}'$ the optimal next period capital stock, that is

$TV(k) = u(F(k) - \tilde{k}') + \beta V(\tilde{k}')$

But now, since $V(k) \leq W(k)$ for all $k \in \mathscr{R}$, we have $K(\tilde{k}') \leq W(\tilde{k}')$, such that it should be clear that

$TV(k) = u(F(k) - \tilde{k}') + \beta V(\tilde{k}')$

$= \max_{k' \in \mathscr{R}} u(F(k) - k') + \beta W(k') = (TW(\tilde{k}'))$ Hence we have shown $V(k) \leq (W)(k)$ that implies $TV(k) \leq (TW)(k)$ and therefore established monotonicity.

III.DISCOUNTING: LET US CONSIDER A CANDIDATE VALUE FUNCTION, *V* , AND A POSITIVE CONSTANT *A*.

$(T(V + a))(k) = \max_{k' \in \mathscr{R}} u(F(k) - k') + \beta V(k') + a)$

$= \max_{k' \in \mathscr{R}} u(F(k) - k') + \beta V(k') + \beta a)$

$= TV(k) + \beta a$

Therefore, the Bellman equation satisfies discounting in the case of optimal growth model.

Hence, the optimal growth model satisfies the Blackwell's sufficient conditions for a contraction mapping, and therefore the value function exists and is unique.

REFERENCES

[1] Sergio Verdu, H.Vincent Poor, “Abstract Dynamic Programme Models under Community Conditions”, Siam J. Control and Optimization, Vol.25, No.4, July 1987.
[2] Wikipedia, “Bellman Equation”, http://en.wikipedia.org/wiki/Bellman_equation
[3] R Bellman, *On the Theory of Dynamic Programming*, Proceedings of the National Academy of Sciences, 1952
[4] S. Dreyfus (2002), 'Richard Bellman on the birth of dynamic programming' *Operations Research* 50 (1), pp. 48–51.