Infinite-Horizon Dynamic Programming

Tianxiao Zheng
SAIF

1. Introduction

Unlike the finite-horizon case, the infinite-horizon model has a stationarity structure in that *both the one-period rewards and the stochastic kernels for the state process are time homogeneous.* Intuitively, we may view the infinite-horizon model as the limit of the finite-horizon model as the time horizon goes to infinity. The difficulty of the infinite-horizon case is that there is no general theory to guarantee the existence of a solution to the Bellman equation. For bounded rewards, we can use the powerful Contraction Mapping Theorem to deal with this issue.

2. Principle of Optimality

The principle of optimality in the infinite horizon states that

1. Dynamic programming principle
   The value function is defined as

$$V(s_t) = \max_{\{a_j\}_{j=t}^{\infty}} \quad \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j u(s_{t+j}, a_{t+j}).$$

   It can be shown that the value function (**independent of time**) satisfies the Bellman equation (a functional equation)

$$V(s) = \max_{a \in \Gamma(s)} \quad u(s,a) + \beta \mathbb{E} V(s'),$$

   where s' is the state variable in the next period, and $\Gamma(s)$ is the set of feasible action a.
2. Verification theorem
   Given any s_0, for any feasible policy π_t, we can use the Bellman equation to derive

$$V^*(s_t) \geq u(s_t, \pi_t) + \beta \mathbb{E}_t V^*(s_{t+1}), \qquad \text{for } t = 0, 1, ..., n$$

   where V^* is the solution to the Bellman equation. Multiplying by β^t and rearranging yield

$$\beta^t V^*(s_t) - \beta^{t+1} \mathbb{E}_t V^*(s_{t+1}) \geq \beta^t u(s_t, \pi_t).$$

Taking expectation conditional on time 0 and summing over $t = 0, 1, ..., n-1$, we obtain

$$V^*(s_0) - \beta^n \mathbb{E}_0 V^*(s_n) \geq \mathbb{E}_0 \sum_{t=0}^{n-1} \beta^t u(s_t, \pi_t)$$

If the **transversality condition** (recall in the finite horizon, $V_T^*(s_T) = u_T(s_T)$)

$$\lim_{n \to \infty} \mathbb{E}_0 \beta^n V^*(s_n) = 0$$

is satisfied, by taking the limit $n \to \infty$ we deduce that

$$V^*(s_0) \geq \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(s_t, \pi_t).$$

The equality holds if we replace π by π^* (optimal policy generated from the Bellman equation). In this case, the right-hand-side of the above inequality becomes $V(s_0)$, giving $V^*(s_0) = V(s_0)$. The result demonstrates that under the transversality condition the solution to the Bellman equation gives the value function for the Markov decision problem. In addition, any plan generated by the optimal policy correspondence from the dynamic programming problem is optimal for the Markov decision problem.

We should emphasize that the transversality condition is a sufficient condition, but not a necessary one. It is quite strong because it requires the limit to converge to zero for any feasible policy. This condition is often violated in many applications with unbounded rewards. However, if any feasible plan that violates the transversality condition is dominated by some feasible plan that satisfies this condition, the solution to the Bellman equation is the value function and that the associated policy function generates an optimal plan. (See Stokey, Lucas, and Prescott, 1989).

3. Any optimal policy obtained from solving the Markov decision problem can be generated by solving the Bellman equation.

3. Solving optimal stopping problem

Here, we give a specific example to show how to solve the optimal stopping problem with the method of dynamic programming.

Example 3.1. *Option exercise*

An agent decides when to exercise an American call option on a stock. Let z_t represents the stock price and I represents the strike price. If the agent chooses to wait, he receives nothing in the current period. Then he draws a new stock price in the next period and decides whether he should exercise the option. In this problem, we formulate the option exercise as a infinite-horizon optimal stopping. Continuation at date t generates a payoff of $f_t(z_t) = 0$, while stopping (option exercise) at date t yields an immediate payoff $g_t(z_t) = z_t - I$ and zero payoff in the future. The decision *maker is risk neutral so that he maximizes his expected return. The discount factor is equal to the inverse of the gross interest rate, $\beta = 1/R$.*

- *Find the Bellman equation of and solve the dynamic programming problem. For simplicity, we assume that z_t is i.i.d. The cumulative distribution function is $F(z)$, $z \in [0, B]$, $B > I$.*
- *Find the mean waiting period until the option is exercised.*

The Bellman equation of the problem

$$\begin{aligned} V(z) &= \max_{a \in \{0,1\}} \quad u(z,a) + \beta \mathbb{E} V(z') \\ &= max(z - I, \beta \mathbb{E} V(z')) \end{aligned} \tag{1}$$

Note that $\beta \mathbb{E} V(z')$ is a constant under iid z_t. Therefore, if $z - I > \beta \mathbb{E} V(z')$, the decision maker chooses to exercise the option and to wait otherwise. As a result,

$$V(z) = \begin{cases} z - I, & \text{if } z > z^*, \\ const., & \text{if } z < z^*. \end{cases}$$

The threshold z^* is determined by $V(z^*) = z^* - I = \beta \mathbb{E} V(z')$,

$$z^* - I = \beta \int_0^{z^*} (z^* - I) dF(z) + \beta \int_{z^*}^{B} (z - I) dF(z) = \beta \int_0^{B} (z^* - I) dF(z) + \beta \int_{z^*}^{B} (z - z^*) dF(z)$$

This gives

$$z^* - I = \frac{\beta}{1 - \beta} \int_{z^*}^{B} (z - z^*) dF(z)$$

From the equation, we see that $z^* \in [I, B]$. The decision maker will not exercise the option for $z_t \in (I, z^*)$ because there is option value of waiting.

The probability of not exercising the option at each period is $\lambda = F(z^*)$. Consequently, the probability of exercising the option at time period t is $\lambda^j (1 - \lambda)$. The mean waiting period is then

$$\sum_{j=0}^{\infty} j \lambda^j (1 - \lambda) = (1 - \lambda) \lambda \frac{d}{d\lambda} \sum_{j=0}^{\infty} \lambda^j = \frac{\lambda}{1 - \lambda}$$

4. Bellman equation as a fixed-point problem

Define a Bellman operator $\hat{T}$ so that

$$\hat{T} f = \max_{a \in \Gamma(s)} u(s, a) + \beta \mathbb{E} f(s'),$$

where f is a continuous function on $\mathbb{S}$. Then the solution to the Bellman equation is a fixed point of $\hat{T}$ in that $\hat{T} V = V$.

The set of bounded and continuous function on the state space $\mathbb{S}$ endowed with the sup norm is a Banach space ($C(\mathbb{S})$). The operator T is a contraction if (1) u is bounded and continuous; (2) Γ is nonempty, compact, and continuous; (3) The stochastic kernel $P(s,a;s')$ satisfies the property that $\int f(s')P(s,a;s')ds'$ is continuous in (s,a) for any bounded and continuous function f; (4) $\beta \in (0,1)$.

The contraction property of the Bellman operator $\hat{T}$ gives the *existence and uniqueness* of the solution to the Bellman equation. It justifies the guess-and-verify method for finding the value function. (As long as we find a solution, it is the solution.) Below is an simple example.

Example 4.1. *A social planner's problem*

$$\max_{\{c_j\}_{j=0}^{\infty}} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \log(c_t)$$

subject to $c_t + K_{t+1} = z_t K_t^{\alpha}$, *where* z_t *follows a Markov process with transition equation* $\log(z_{t+1}) = \rho \log(z_t) + \sigma \varepsilon_{t+1}$. *Here,* $\rho \in (0,1)$ *and* ε_t *is normal distribution with mean 0 and variance 1.*

We write the Bellman equation as

$$V(K,z) = \max_c \ \log c + \beta \mathbb{E} V(zK^{\alpha} - c, z')$$

Given the log utility, we could guess the value function takes the functional form $V(K,z) = d_0 + d_1 \log z + d_2 \log K$. *The maximization problem*

$$\max_c \ \log c + \beta \mathbb{E} V(zK^{\alpha} - c, z') = \max_c \ \log c + \beta d_0 + \beta d_2 \log(zK^{\alpha} - c) + \beta d_1 \mathbb{E} \log z'$$

yields $c = zK^{\alpha}/(1+\beta d_2)$. *Therefore,*

$$d_0 + d_1 \log z + d_2 \log K = \log \frac{zK^{\alpha}}{1+\beta d_2} + \beta d_2 \log \frac{zK^{\alpha}\beta d_2}{1+\beta d_2} + \beta d_1 \rho \log z + \beta d_0$$

Comparing the coefficients on both side yields

$$\begin{aligned} d_2 &= \alpha + \alpha\beta d_2, \\ d_1 &= 1 + \beta d_2 + \beta\rho d_1, \\ d_0 &= -\log(1+\beta d_2) + \beta d_2 \log \frac{\beta d_2}{1+\beta d_2} + \beta d_0. \end{aligned} \tag{2}$$

Solving these equations, we have

$$\begin{aligned} d_2 &= \frac{\alpha}{1-\alpha\beta}, \\ d_1 &= \frac{1}{(1-\alpha\beta)(1-\rho\beta)}, \\ d_0 &= \frac{1}{1-\beta}[\log(1-\alpha\beta) + \frac{\alpha\beta}{1-\alpha\beta}\log(\alpha\beta)] \end{aligned} \tag{3}$$

The decision rule can also be derived: $K_{t+1} = \alpha\beta z_t K_t^{\alpha}$.

4.1. Value function iteration

The contraction property of the Bellman operator $\hat{T}$ also gives a *globally* convergent algorithm to solve for the value function. Specifically, for any $v_0 \in C(\mathbb{S})$, because $\hat{T}$ is a contraction operator

$$\lim_{N\to\infty} \hat{T}^N v_0 = V.$$

This property gives rise to a numerical algorithm known as *value function iteration* for finding V.

We start with an arbitrary guess $V_0(s)$ and iterate the Bellman operator

$$V_1 = \hat{T}V_0, \;\; V_2 = \hat{T}V_1, \; ... \; V_n(s) = \hat{T}V_{n-1} = \max_{a\in\Gamma(s)} u(s,a) + \beta\mathbb{E}V_{n-1}(s'), ...$$

until V_n is convergent. The contraction mapping theorem guarantees the convergence of this algorithm. In particular, the contraction property implies that $\|V_n(s) - V(s)\|$ converges to zero at a geometric rate.

Note that in the case where we have set $v_0(s) = 0$, the value function iteration algorithm is equivalent to solving a finite horizon problem by backward induction. Suppose we stop the iteration at N because convergence is attained, e.g. $\|V_N(s) - V_{N-1}(s)\| \approx 10^{-15}$. The equivalent finite horizon problem is then define with $u_t(s_t, a_t) = u(s_t, a_t)$, for $t = 0, 1, ..., N-1$ and $u_N(s_N, a_N) = 0$.

4.2. Policy function iteration (Howard's improvement algorithm)

We digress here to introduce a usually much faster algorithm to solve the Bellman equation. It is known as policy function iteration and consists of the following three steps:

1. Choose a arbitrary policy g_0, and compute the value function associated implied by g_0.

   $$V_0(s) = u(s, g_0(s)) + \beta\mathbb{E}V_0(s').$$

   On discretized grids of the state space, this is usually done by solving a linear system. There also exists a fast method to compute $V_0(s)$ by defining an operator $\hat{B}$,

   $$\hat{B}V_0(s) = u(s, g_0(s)) + \beta\mathbb{E}V_0(s'),$$

   and finding the fix point $V_0 = \hat{B}V_0$. Iterate on B for a small number of times to obtain an approximation of V_0.
2. Generate a improved policy $g_1(s)$ that solves the two-period problem

   $$\max_{a\in\Gamma(s)} \quad u(s,a) + \beta\mathbb{E}V_0(s')$$

3. Given g_1, one continues the cycle of value function evaluation step and the policy improvement step until the first iteration n such that $\|g_n - g_{n-1}\| \to 0$ (or alternatively $\|V_n - V_{n-1}\| \to 0$). Since such a g_n satisfies the Bellman equation, it is optimal.

5. Application to optimal control

The associated dynamic programming problem of infinite-horizon optimal control is give by

$$V(x,z) = \max_{a \in \Gamma(x,z)} u(x,z,a) + \beta \int V(\phi(x,a,z,z'),z')Q(z,z')dz'$$

5.1. Characterization of the value function: monotonicity, concavity and differentiability

Analogous to the finite-horizon optimal control, we have the following properties of the solution to the Bellman equation (*i.e.* V and policy G):

- under the condition (1) $u(.,z,a)$ is continuous and bounded for each z, a, (2) $u(.,z,a)$ is strictly increasing, (3) for each z, $\Gamma(.,z)$ is increasing ($x < x'$ implies $\Gamma(x,z) \subset \Gamma(y,z)$), (4) $\phi(.,a,z,z')$ is increasing for each a, z, z', *then* $V(.,z)$ *is strictly increasing for each* z.
- under the condition (1), (2), (3), (4), (5) at each z, for all x, a, x', a' and $\theta \in (0,1)$, $u(\theta x + (1-\theta)x', z, \theta a + (1-\theta)a') \geq \theta u(x,z,a) + (1-\theta)u(x',z,a')$, (6) $\phi(.,.,z,z')$ is concave for each z, z', *then* $V(.,z)$ *is strictly concave for each* z*;* G *is a single-valued continuous function.*
- under (5), (7) for each z, $u(.,z,.)$ is continuously differentiable on the interior of $\mathbb{X} \times \mathbb{A}$, (8) for each z, z', $\phi(.,.,z,z')$ is differentiable on the interior of $\mathbb{X} \times \mathbb{A}$, (9) at each z, for all x, x', $a \in \Gamma(x,z)$ and $a' \in \Gamma(x',z)$ imply that $\theta a + (1-\theta)a' \in \Gamma(\theta x + (1-\theta)x', z)$, *then* $V(.,z)$$ *is continuously differentiable.*

The **envelope condition** is then given by

$$V_x(x,z) = u_x(x,z,a) + \beta \int V_x(\phi(x,a,z,z'),z')\phi_x(x,a,z,z')Q(z,z')dz'$$

The **first order condition** is given by

$$0 = u_a(x,z,a) + \beta \int V_x(\phi(x,a,z,z'),z')\phi_a(x,a,z,z')Q(z,z')dz'$$

5.2. Maximum Principle

In this note, we do not discuss problems with *unbounded rewards*. The difficulty is that there is no general fixed-point theorem to guarantee the existence of a solution. For readers interested in this problem, we pointed out two relevant work. Alvarez and Stokey, 1998 consider general dynamic programming problems with a homogeneity property. They show that the Weighted Contraction Mapping Theorem can be applied for general cases with positive degree of homogeneity. Durán, 2000 extends the weighted norm approach to general problems without the homogeneity property.

Analyzing the existence and properties of the value function is nontrivial for unbounded reward functions. By contrast, unbounded reward functions do not pose any difficulty for the Maximum Principle to work. To present the infinite horizon maximum principle, we write the Lagrangian form for the optimal control problem

$$\mathcal{L} = E\left[\sum_{t=0}^{\infty} \beta^t u(x_t, z_t, a_t) - \beta^{t+1}\mu_{t+1}(x_{t+1} - \phi(x_t, z_t, a_t, z_{t+1}))\right]$$

F.O.C

$$\begin{aligned} a_t: &\quad 0 = u_a(x_t, z_t, a_t) + \beta \mathbb{E}_t \mu_{t+1} \phi_a(x_t, z_t, a_t, z_{t+1}), \quad t > 0 \\ x_t: &\quad \mu_t = u_x(x_t, z_t, a_t) + \beta \mathbb{E}_t \mu_{t+1} \phi_x(x_t, a_t, z_t, z_{t+1}), \quad t \geq 1 \end{aligned} \tag{4}$$

Setting $\mu_t = V_x(x_t, z_t)$, we can see that the two conditions above are equivalent to the first order condition and envelope condition of Bellman equation. The Lagrange multiplier μ_t is interpreted as the shadow value of the value function.

Recall that in the finite horizon case, we have a terminal condition $\mu_T = \frac{\partial u_T}{\partial x_T}$ to solve the problem by backward induction. There is no well-defined terminal condition in the infinite horizon case. Here, a sufficient boundary condition is in the form of transversality condition

$$\lim_{T \to \infty} \mathbb{E}[\beta^T \mu_T x_T] = 0$$

For a special class of control problems -- the Euler class, we could prove the transversality condition is also necessary. (cf. Ekeland and Scheinkman, 1986 and Kamihigashi, 2000)

5.2.1. Euler class

In practice, it may be possible to use simple tricks to transform the general optimal control problem to a special class of control problems -- the Euler class. Suppose it is possible to perform a change of variables such that the state transition equation becomes

$$x_{t+1} = a_t.$$

This could simplify the solution by Bellman equation and maximum principle.

1. Bellman equation

   The envelope condition becomes very simple

   $$V_x(x, z) = u_x(x, z, g(x, z))$$

   where $x' = a = g(x, z)$ is the optimal policy. Substituting the envelope condition into the first order condition yields the Euler equation

$$\begin{aligned}0 &= u_a(x,z,a) + \beta \int u_x(x',z',a')Q(z,z')dz' \\ &= u_a(x,z,g(x,z)) + \beta \int u_x\Big(g(x,z), z', g\big(g(x,z)\big)\Big)Q(z,z')dz'\end{aligned}$$

This is a functional equation for the optimal policy g. Instead of solving the original Bellman equation, for the Euler class, we could solve the Euler equation.

2. Maximum principle

The first order conditions become

$$\begin{aligned}a_t: &\quad 0 = u_a(x_t,z_t,a_t) + \beta\mathbb{E}_t[\mu_{t+1}], \quad t > 0 \\ x_t: &\quad \mu_t = u_x(x_t,z_t,a_t), \quad t \geq 1\end{aligned} \tag{5}$$

Substituting the second equation into the first one, we get the sequential form of Euler equation

$$0 = u_a(x_t,z_t,x_{t+1}) + \beta\mathbb{E}_t[u_x(x_{t+1},z_{t+1},x_{t+2})].$$

The transversality condition can be expressed as

$$\lim_{T\to\infty} \mathbb{E}[\beta^T \mu_T x_T] = \lim_{T\to\infty} \mathbb{E}[\beta^T u_x(x_T,z_T,x_{T+1})x_T] = 0.$$

By using the Euler equation, we have

$$\begin{aligned}\lim_{T\to\infty} \mathbb{E}[\beta^T u_x(x_T,z_T,x_{T+1})x_T] &= \lim_{T\to\infty} \mathbb{E}[\beta^{T-1} a_{T-1}\beta\mathbb{E}_{T-1}u_x(x_T,z_T,x_{T+1})] \\ &= -\lim_{T\to\infty} \mathbb{E}[\beta^{T-1}u_a(x_{T-1},z_{T-1},x_T)a_{T-1}].\end{aligned} \tag{6}$$

Therefore, the condition can be rewritten as

$$\lim_{T\to\infty} \mathbb{E}[\beta^T u_a(x_T,z_T,x_{T+1})a_T] = \lim_{T\to\infty} \mathbb{E}[\beta^T u_a(x_T,z_T,x_{T+1})x_{T+1}] = 0.$$

To get some economic sense of the transversality condition, we consider a simple example.

Example 5.1. *A social planner's resource allocation problem.*
The planner's objective is to choose sequences of consumption (c_t) so as to

$$\max_{\{c_t\}_{t=0}^T} \quad \mathbb{E}\left[\sum_{t=0}^T \beta^t u(c_t)\right], \quad \beta \in (0,1)$$

subject to the resource constraint

$$K_{t+1} = z_t F(K_t) - c_t, \quad (x_0, z_0) \text{ given.}$$

By defining $K_{t+1} \equiv a_t$, the problem can be rewritten as

$$\max_{\{c_t\}_{t=0}^{T}} \quad \mathbb{E}\left[\sum_{t=0}^{T} \beta^t u(z_t F(K_t) - a_t)\right], \quad \beta \in (0,1)$$

subject to $K_{t+1} = a_t$. (K_t is the state variable; a_t is the control variable.)
In the last period, the agent solves the problem

$$\max_{K_{T+1}} \quad \mathbb{E}[\beta^T u(z_T F(K_T) - K_{T+1})]$$

K_{T+1} should be non-negative.

- *If $K_{T+1} = 0$, the following condition should be satisfied $\mathbb{E}[\beta^T u'(c_T)] > 0$;*
- *If $K_{T+1} > 0$, the following condition should be satisfied $\mathbb{E}[\beta^T u'(c_T)] = 0$*

We can combine the conditions as $\mathbb{E}[\beta^T u'(c_T) K_{T+1}] = 0$. This is the transversality condition in the finite horizon case. The economic meaning is that the expected discounted shadow value of the terminal state (e.g., capital or wealth) must be zero. In the infinite horizon case, we take the limit of the condition.

Finally, let us consider an example to illustrate the above theoretical results.

Example 5.2. *A consumption-saving problem*

$$\max_{\{c_t\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{T} \frac{c_t^{1-\gamma}}{1-\gamma}\right], \quad \gamma > 0, \gamma \neq 1$$

subject to $x_{t+1} = R_{t+1}(x_t - c_t)$, $x_{t+1} > 0$, $x_0 > 0$ given, where $R_{t+1} > 0$ is i.i.d. drawn from a distribution. By defining $y_{t+1} = x_{t+1}/R_{t+1}$, $y_{t+1} = a_t = x_t - c_t = y_t R_t - c_t$, $c_t = y_t R_t - a_t$.

F.O.C

$$\begin{aligned} a_t: &\quad 0 = u_a(x_t, z_t, a_t) + \beta \mathbb{E}_t \mu_{t+1} = -(y_t R_t - a_t)^{-\gamma} + \beta \mathbb{E}_t \mu_{t+1} = -c_t^{-\gamma} + \beta \mathbb{E}_t \mu_{t+1} \\ x_t: &\quad \mu_t = u_x(x_t, z_t, a_t) = R_t (y_t R_t - a_t)^{-\gamma} = R_t c_t^{-\gamma}, \end{aligned} \tag{7}$$

The resulting Euler equation is

$$c_t^{-\gamma} = \beta \mathbb{E}_t[R_{t+1} c_{t+1}^{-\gamma}]$$

An obvious guess of the consumption policy is that $c_t = \mathcal{C} x_t$ ($0 < \mathcal{C} < 1$). Plugging the conjecture into the Euler equation yields

$$(\mathcal{C} x_t)^{-\gamma} = \beta \mathbb{E}_t[R_{t+1} (\mathcal{C} x_{t+1})^{-\gamma}] = \beta \mathbb{E}_t[R_{t+1} \mathcal{C}^{-\gamma} (R_{t+1} x_t - R_{t+1} \mathcal{C} x_t)^{-\gamma}]$$

The above equation gives us $\mathcal{C} = 1 - (\beta \mathbb{E}_t[R_{t+1}^{1-\gamma}])^{1/\gamma}$. *Consider the Bellman equation*

$$\begin{aligned} V(x) &= \max_c \quad u(c) + \beta \mathbb{E} V(x'), \\ &= u(\mathcal{C}x) + \beta \mathbb{E} V(R'x(1-\mathcal{C})) = \frac{(\mathcal{C}x)^{1-\gamma}}{1-\gamma} + \beta \mathbb{E} V(R'x(1-\mathcal{C})) \end{aligned} \tag{8}$$

An obvious guess of the value function is $V(x) = \mathcal{B}x^{1-\gamma}/(1-\gamma)$. *Plugging the conjecture into the Euler equation yields*

$$\mathcal{B} = \left[1 - (\beta \mathbb{E}_t[R_{t+1}^{1-\gamma}])^{1/\gamma}\right]^{-\gamma}$$

Now, we could check the transversality condition

$$\lim_{t\to\infty} \mathbb{E}_0 \beta^t V(x_t) = 0$$

$$\begin{aligned} \mathbb{E}_0 \beta^t V(x_t) &= \frac{\beta^t \mathcal{B}}{1-\gamma} \mathbb{E}_0[x_t^{1-\gamma}] = \frac{\beta^t \mathcal{B}}{1-\gamma} \mathbb{E}_0[R_t^{1-\gamma}(x_{t-1} - c_{t-1})^{1-\gamma}] \\ &= \frac{\beta^t \mathcal{B}}{1-\gamma}(1-\mathcal{C})^{1-\gamma} \mathbb{E}_0[R_t^{1-\gamma} x_{t-1}^{1-\gamma}] \\ &= \frac{\beta^t \mathcal{B}}{1-\gamma}(1-\mathcal{C})^{t(1-\gamma)} x_0^{1-\gamma} \mathbb{E}_0[\prod_{j=1}^{t} R_j^{1-\gamma}] \end{aligned} \tag{9}$$

and the transversality condition

$$\lim_{T\to\infty} \mathbb{E}_0[\beta^T u_x(x_T, z_T, x_{T+1}) x_T] = \lim_{T\to\infty} \mathbb{E}_0[\beta^T R_T c_T^{-\gamma} x_T / R_T] = \lim_{T\to\infty} \mathbb{E}_0[\beta^T c_T^{-\gamma} x_T]$$

$$\mathbb{E}_0[\beta^T c_T^{-\gamma} x_T] = \beta^T \mathcal{B} \mathbb{E}_0[x_T^{1-\gamma}]$$

References

Alvarez, F., Stokey, N. L., 1998. Dynamic programming with homogeneous functions. Journal of Economic Theory 82, 167 – 189.

Durán, J., 2000. On dynamic programming with unbounded returns. Economic Theory 15, 339–352.

Ekeland, I., Scheinkman, J. A., 1986. Transversality conditions for some infinite horizon discrete time optimization problems. Mathematics of Operations Research 11, 216–229.

Kamihigashi, T., 2000. A simple proof of ekeland and scheinkman's result on the necessity of a transversality condition. Economic Theory 15, 463–468.

Stokey, N. L., Lucas, R. E., Prescott, E. C., 1989. Recursive Methods in Economic Dynamics. Harvard University Press.