Continuous Finite Horizon Markov Decision Processes

Discrete Approximation Methods

- Discrete Approximation methods compute the value function V_t and decision rules δ_t, for $t = 1, \cdots, T$, at a finite set of points in the state space S and constraint sets $A(s)$. We can refer to this finite subset of points of S as a grid, and denote it by $s_1, \cdots, s_N$. This definition does not presume evenly spaced points.

- One first chooses grids for the state and action spaces, and then performs backward induction for a finite MDP problem defined on these grids. Thus discretization amounts to replacing the continuous-state version of the backward induction step

$$V_t(s) = \Gamma(V_{t+1})(s) = \max_{a \in A(s)} \left[u(s,a) + \beta \int V_{t+1}(s') p(ds'|s,a) \right], \ s \in S, \tag{1}$$

  by a discretized equivalent

$$\hat{V}_t(s) = \hat{\Gamma}(\hat{V}_{t+1})(s) = \max_{a \in \hat{A}(s)} \left[u(s,a) + \beta \sum_{k=1}^{N} \hat{V}_{t+1}(s_k) p_N(s_k|s,a) \right], \ s \in s_1, \cdots, s_N. \tag{2}$$

- There is no particular reason to assume that the number of points N making up the grid of S is the same as the number of grid points making up the grids of the constraint sets $A(s)$. It is even not necessary to discretize the constraint set if continuous optimization methods are used.

- The key restriction is that the state space is discretized, which implies that discrete approximation methods amount to a replacement of the true Bellman operator $\Gamma : B(S) \longrightarrow B(S)$ (where $B(S)$ denotes the infinite dimensional Banach space of measurable, bounded functions), by an approximate "discretized" Bellman operator $\hat{\Gamma}_N : R^N \longrightarrow R^N$, but otherwise the backward induction process is identical.

- A Banach space is a *complete* normed space, meaning that every Cauchy sequence converges. Completeness is an important technical property for a space, since often we may be able to prove that a sequence satisfies the Cauchy criterion and we want to deduce that it converges to some element of the space. Intuitively a Cauchy sequence means that a sequence x_n of real numbers *bunches up*, that is, all the elements of the sequence are arbitrarily close to one another sufficiently far out in the sequence.

- However, even though the discrete solution $\hat{V} = (\hat{V}_1, \cdots, \hat{V}_T)$ consists of T vectors in R^N, we can use these vectors as data for a variety of interpolation or "smoothing" methods to produce continuous solutions defined over the entire state space S.

- In order to solve the problem we need to choose the N grid points $s_1, \cdots, s_N$, at which to evaluate V_{t+1}, and how to construct a discretized estimate p_N of the continuous transition probability p. The latter usually follows from the former. At this time we will focus on using uniform and quadrature grids to construct the grid and solve this problem.

Uniform Grids

- It is the most obvious way to discretize a continuous state space, consisting of equi-spaced points in S and A. Specifically, we partition the d_s-dimensional cube $[0,1]^{d_s}$ into equal subcubes of length h on each side. This results in a total of $N = (1/h)^{d_s}$ subcubes. The amount of work increases exponentially in the dimensions of d_s and d_a.
- Let S_h denote the partition of S induced by the uniform partition of the unit cube. Similarly, denote A_h as the partition induced by the discretization of the action space A. S_h consists of $N_s = (1/h)^{d_s}$ equal subcubes of length h on each side, and similarly A_h. Let s_k denote an arbitrary element (grid point) in the k_{th} partition element of S_h, and let $k(s)$ denote the index of the partition element of S_h that contains a given point $s \in S$. And in a similar way denote a_k.
- The discretized utility and transition probability are defined by:

$$u_h(s,a) = u(s_{k(s)}, a), \tag{3}$$

$$p_h(s'|s,a) = \frac{p(s_{k(s')}|s_{k(s)}, a)}{\int p(s_{k(s')}|s_{k(s)}, a)ds'}, \tag{4}$$

  where the normalization of the second equation insures that p_h is a well defined transition probability density on S. In most cases we will need to do numerical integration to compute the normalizing factors in the denominator. This integration can be carried out by any method, for example quadrature.
- Given these objects we can define the discretized Bellman operator by

$$\hat{\Gamma}_h(V)(s) = \max_{a_k, k=1,\cdots,a_{N_a}} \left[u_h(s,a_k) + \beta/N \sum_{k=1}^{N} V(s_k) p_h(s_k|s_{k(s)}, a_k) \right]. \tag{5}$$