Definition of the *state* in dynamic programming

Consider a dynamic-programming problem in finite time: $t = 0, 1, \dots, T$. Every period, the agent chooses a control $u_t \in \mathcal{R}^n$ from a feasible set. This feasible set may depend on past policies $\{u_s\}_{s=0}^t$, past shocks $\{\epsilon_s\}_{s=0}^t$ and parameters[1] θ of the model:

$$u_t \in \Phi_t\Big(\{u_r\}_{r=0}^s, \{\epsilon_r\}_{r=0}^s; \theta\Big),$$

where $\{\Phi_t\}_{t=0}^T$ is a sequence of correspondences.

The agent's per-period return function $G_t(\cdot)$, just as the feasible set, may depend on all past actions, past shocks and the parameters:

$$G_t\Big(\{u_s\}_{s=0}^t, \{\epsilon_s\}_{s=0}^t; \theta\Big)$$

The agent's total payoff (or value) at time t is assumed to be additive in the per-period returns. We also usually assume that the future is discounted at the constant rate β:[2]

$$E_t \sum_{s=t}^{T} \beta^{s-t} G_s\Big(\{u_r\}_{r=0}^s, \{\epsilon_r\}_{r=0}^s; \theta\Big), \tag{1}$$

where $\{G_t(\cdot)\}_{t=0}^T$ is a sequence of real-valued functions.

Shortly speaking, the state is the *smallest set of variables that summarize the effect of past actions and past shocks on the current physical environment.* Note that we do *not* include parameters of the model into the state.

Definition: Formally, the *state* x_t of a problem is defined as the *smallest* set of variables which, given the parameters of the model, determine the following:

- The set of feasible policies $\{u_s\}_{s=t}^T$
- The payoff at t given any future policy, see (1).

So very often, a shock will form part of the state, though usually not all of its past realizations. In the deterministic case, it is usually possible to control the state x' tomorrow perfectly, so it is natural to use it directly

[1] Parameters are (deterministic) variables that cannot be influenced by the agent's actions.

[2] Exercise: Is the latter strictly necessary for dynamic programming to be valid in the finite-horizon case?

as the control variable (as we did in the neo-classical growth model and the consumption-savings problem). As a convention, we also always include time t into the state in a finite-horizon problem since the payoff (1) depends on the length of the time horizon.

Together with the state, we always want to find the feasibility correspondence, $u_t \in \Gamma_t(x_t)$, and the per-period return function $F_t(x_t, u_t)$. These allow us to write down the Bellman equation:

$$\begin{aligned} V_t(x) &= \max_{u \in \Gamma_t(x)} \Big\{ F_t(x, u) + \beta E_t\big[V_{t+1}(x')\big] \Big\} \\ \text{s.t.} \quad x' &= H(x, u, \epsilon'), \end{aligned}$$

where $H(\cdot)$ is the *law of motion*, which tells us the state tomorrow as a function of the state x today, the control/action today u and the shock ϵ' tomorrow. Note that also the conditional probabilities implicit in the conditional-expectations operator often depend on the state.

Example: Deterministic consumption-savings problem

In the deterministic consumption-savings problem with a no-borrowing constraint, the objects are classified as follows:

- Parameters: β, R, $u(\cdot)$, a_0, the deterministic sequence $\{w_t\}_{t=0}^T$.
- State: t, a_t.
- Shocks: None.
- Control: a_{t+1}
- Feasible set: $\Gamma_t(a) = [0, R(w_t + a)]$.
- Return function: $F_t(a_t, a_{t+1}) = u(w_t + a_t - a_{t+1}/R)$
- Payoff at t: $\sum_{s=t}^T \beta^{s-t} u(w_t + a_t - a_{t+1}/R)$.
- Law of motion: Trivial, the state tomorrow is the control today.

The Bellman equation is

$$V_t(a) = \max_{a' \in [0, R(w_t + a)]} \Big\{ u(w_t + a - a'/R) + \beta V_{t+1}(a') \Big\}.$$