Example – Lagged Inventory

A modified store inventory management example. The orders u_k are ready with a one-day lag (think of it as a production/delivery time). Additionally, there is no backlog option for orders. One state variable x_k will describe the amount of the product available in the morning (when the store opens). During the day, the demand w_k is realized. In the evening (after the store closes), the order from the previous day arrives and is prepared for the next day.

The system develops through the following equation:

$$x_{k+1} = \max\{x_k - w_k, 0\} + u_{k-1}.$$

In order to apply DP, we need to augment this system – introducing another state variable $s_k = u_{k-1}$, the system equation have the form:

$$x_{k+1} = \max\{x_k - w_k, 0\} + s_k,$$
$$s_{k+1} = u_k.$$

An additional modification will encode the maximum size of the inventory M – any access over that M that is stored over night is simply lost (more complicated variations require a more complex model). The new augmented system has the following form:

$$x_{k+1} = \min\{\max\{x_k - w_k, 0\} + s_k, M\}$$
$$s_{k+1} = u_k.$$

The last restriction is w.r.t. the amount ordered: $0 \le u_k \le U$.

The costs will encompass the cost of the inventory c_1, the order cost c_2, and the profit from selling the items c_3 (negative):

$$g(x_k, s_k, u_k, w_k) = c_1(x_k) + c_2(u_k) + c_3(x_k, w_k),$$

and the terminal cost will be just the cost of the inventory

$$g(x_N) = c_1(x_N).$$

A big advantage of DP is that the cost functions can be rather arbitrary (no need for linearity, convexity, or even continuity), e.g.

$$c_1(x) = \sqrt{x},$$
$$c_2(u) = \begin{cases} 0, & \text{if } u = 0. \\ 0.2 + u, & \text{otherwise.} \end{cases}$$
$$c_3(x, w) = -3 \cdot \min\{x_k, w_k\}.$$

Assignment: Compute the optimal policy (minimizing the expected costs) for

$$N = 15, \quad M = 4, \quad u_k \in [0, \dots, 3], \quad w_k = \begin{cases} 0, & p(0) = 0.1 \\ 1, & p(1) = 0.3 \\ 2, & p(2) = 0.5 \\ 3, & p(3) = 0.1 \end{cases}$$