Battery Storage Management System through Reinforcement Learning

Markov Decision Process (MDP):

Set of States ($\mathcal{S}$):

$$\mathcal{S} = \{s_t \mid s_t = (d_t, \mathrm{SOC}_t, p_t)\}$$

d_t : electricity demand at time t
SOC_t : state of charge of the battery at time t
p_t : price of grid electricity at time t

Set of Actions ($\mathcal{A}$):

$$\mathcal{A} = \{a_t \mid a_t = b_t\}$$

If $b_t > 0$: charge mode, else: discharge mode. b_t can be chosen to be continuous or discrete.

Note that the energy supplied by the grid is:

$$g_t = d_t + b_t$$

Assume that every energy not met by the battery is met by the grid; the above equality holds for charge and discharge.

Transition Dynamics:

$$\mathrm{SOC}_{t+1} = \mathrm{SOC}_t + b_t$$

Reward Function:

$$r_t = \begin{cases} (p_t(d_t + b_t))^2 + \lambda(b_{t+1} - b_t)^2 & \text{if } \mathrm{SOC}_{\min} \leq \mathrm{SOC}_t \leq \mathrm{SOC}_{\max} \\ \text{Penalty} & \text{otherwise} \end{cases}$$

Discount Factor:

$\gamma = 1$