Introduction to Mathematical Finance

Exercise sheet 11

Exercise 11.1 Consider a financial market with positive prices and i.i.d returns: $X_0^l > 0$ and $X_{k+1}^l = X_k^l R_{k+1}^l$, $l = 1, \dots, d$, $k = 0, \dots, T-1$, where $R_1, \dots R_T$ are i.i.d with values in $\mathbb{R}_{++}^d$. Suppose that the filtration is given by the natural filtration of the stochastic process R, i.e. $\mathbb{F} = (\mathcal{F}_k)_k$ where $\mathcal{F}_k = \sigma(R_j, j \leq k)$.
Consider the utility maximization problem with objective function

$$\mathcal{U}(\vartheta, c) = E\left[\sum_{j=0}^{T-1} U_c(c_j) + U_w(W_T^{v_0,\vartheta,c})\right]$$

As in Section III.1, we use the parameterization of the above optimization problem in terms of the proportion of time k wealth invested in asset l (denoted π_k^l) and the proportion of time k wealth spent on consumption (denoted γ_k):

$$\begin{aligned} W_{k+1}^{v_0,\pi,\gamma} &= W_k^{v_0,\pi,\gamma}\left(1 + \pi_k \cdot (R_{k+1} - 1) - \gamma_k\right) \\ c_k &= \gamma_k W_k^{v_0,\pi,\gamma} \\ \vartheta_{k+1}^l &= \frac{\pi_k^l W_k^{v_0,\pi,\gamma}}{X_k^l} \end{aligned}$$

Recall that $c_T = 0$; so $\gamma_T = 0$.
We thus consider the optimization problem

$$\mathcal{U}(\pi, \gamma) = E\left[\sum_{j=0}^{T-1} U_c(c_j) + U_w(W_T^{v_0,\pi,\gamma})\right] \tag{1}$$

In the lecture we solved (1) in the case when both the consumption and the final wealth utility functions were given by a power utility: $U_c(x) = U_w(x) = x^\alpha$ for some $\alpha > 0$. In this exercise you are asked to solve (1) in the case when the utility functions are given by the log utility: $U_c(x) = U_w(x) = \log(x)$.
Hint: Use Dynamic Programming Principle

Solution 11.1 The dynamic programming principle (Proposition III.2.5) indicates that we should solve

$$\sigma_k(v_k) = \operatorname{ess\,sup}_{\pi_k,\gamma_k}\left(U_c(\gamma_k v_k) + E\left[\sigma_{k+1}\left(v_k(1 + \pi_k \cdot (R_{k+1} - 1) - \gamma_k)\right) | \mathcal{F}_k\right]\right)$$

by backwards induction to get candidate optimal solution.

- First note that $\sigma_T(v_T) = U_w(v_T) = \log(v_T)$

- Consider $k = T-1$:

$$
\begin{aligned}
\sigma_{T-1}(v_{T-1}) &= \operatorname{ess\,sup}_{\pi_{T-1},\gamma_{T-1}} \Big(\log(\gamma_{T-1}v_{T-1}) + E\left[\sigma_T\left(v_{T-1}(1+\pi_{T-1}\cdot(R_T-1)-\gamma_{T-1})\right)|\mathcal{F}_{T-1}\right]\Big)\\
&= \operatorname{ess\,sup}_{\pi_{T-1},\gamma_{T-1}} \Big(\log(\gamma_{T-1}) + \log(v_{T-1}) + \\
&\qquad E\left[\log(v_{T-1}) + \log\left(1+\pi_{T-1}\cdot(R_T-1)-\gamma_{T-1}\right)|\mathcal{F}_{T-1}\right]\Big)\\
&= 2\log(v_{T-1}) + \operatorname{ess\,sup}_{\pi_{T-1},\gamma_{T-1}} \Big(\log(\gamma_{T-1}) + E\left[\log((1+\pi_{T-1}\cdot(R_T-1)-\gamma_{T-1})\,|\mathcal{F}_{T-1}\right]\Big)\\
&= 2\log(v_{T-1}) + \\
&\qquad \operatorname{ess\,sup}_{\pi_{T-1},\gamma_{T-1}} \Big(\log(\gamma_{T-1}) + E\left[\log((1+\pi\cdot(R_T-1)-\gamma)\right]_{|\pi=\pi_{T-1},\gamma=\gamma_{T-1}}\Big)\\
&= 2\log(v_{T-1}) + C_{T-1}
\end{aligned}
$$

where

$$
C_{T-1} := \sup_{\pi,\gamma}\Big(\log(\gamma) + E\left[\log(1+\pi\cdot(R-1)-\gamma)\right]\Big)
$$

In the fourth equality above, we have used that R_T is independent of $\mathcal{F}_{T-1}$ and π_{T-1} and γ_{T-1} are $\mathcal{F}_{T-1}$ measurable. The random variable R appearing in the definition of C_{T-1} is distributed identically as the i.i.d random variables R_k.

- For a general $0 \le k \le T-1$, if $\sigma_{k+1}(v_{k+1}) = (T-(k+1)+1)\log(v_{k+1}) + \sum_{j=k+1}^{T-1} C_j$ for some constants C_j, then:

$$
\begin{aligned}
\sigma_k(v_k) &= \operatorname{ess\,sup}_{\pi_k,\gamma_k}\Big(\log(\gamma_k v_k) + E\left[\sigma_{k+1}\left(v_k(1+\pi_k\cdot(R_{k+1}-1)-\gamma_k)\right)|\mathcal{F}_k\right]\Big)\\
&= \operatorname{ess\,sup}_{\pi_k,\gamma_k}\Big(\log(\gamma_k) + \log(v_k) + \\
&\qquad E\left[(T-(k+1)+1)\{\log(v_k) + \log\left(1+\pi_k\cdot(R_{k+1}-1)-\gamma_k\right)\} + \sum_{j=k+1}^{T-1} C_j|\mathcal{F}_k\right]\Big)\\
&= (T-k+1)\log(v_k) + \sum_{j=k+1}^{T-1} C_j\\
&\qquad + \operatorname{ess\,sup}_{\pi_k,\gamma_k}\Big(\log(\gamma_k) + (T-k)E\left[\log((1+\pi_k\cdot(R_{k+1}-1)-\gamma_k)\,|\mathcal{F}_k\right]\Big)\\
&= (T-k+1)\log(v_k) + \sum_{j=k+1}^{T-1} C_j\\
&\qquad \operatorname{ess\,sup}_{\pi_k,\gamma_k}\Big(\log(\gamma_k) + (T-k)E\left[\log((1+\pi\cdot(R_{k+1}-1)-\gamma)\right]_{|\pi=\pi_k,\gamma=\gamma_k}\Big)\\
&= (T-k+1)\log(v_k) + \sum_{j=k}^{T-1} C_j
\end{aligned}
$$

where

$$C_k := \sup_{\pi,\gamma} \Big(\log(\gamma) + (T-k) E\left[\log(1 + \pi \cdot (R-1) - \gamma)\right] \Big)$$

This proves by induction that for all $0 \leq k \leq T-1$,

$$\sigma_k(v_k) = (T-k+1)\log(v_k) + \sum_{j=k}^{T-1} C_j$$

Assume the constants C_k are finite and that there exist maximisers (π_k^*, γ_k^*) (constrained to statisfy $\gamma_k^* \geq 0$ and $1^T \pi_k^* = 1$ for all k). We then expect by the dynamic programming principle (π_k^*, γ_k^*) to be optimal for the original static utility optimization problem (1). Indeed the dynamic programming principle (Proposition III.2.5) tells us that for an optimal solution of (1), $\sigma_k(v_k)$ computed above is equal to the optimal remaining utility at time k. In particular,

$$E[\sigma_0(v_0)] = \mathcal{U}^*$$

is the optimal value of (1). To find the optimal strategy achieving this value $\mathcal{U}^*$, we first calculate backwards $\sigma_k(v_k)$ and the corresponding optimizers $\pi_k^*(v_k)$ and $\gamma_k^*(v_k)$. In general this backward step gives the optimal strategy *as a function of the current wealth* v_k. To get the optimal strategy we then need to iterate forward starting from the initial wealth v_0: $W_0 = v_0$ and if W_k is known, set

$$\begin{aligned}
\pi_k^* &= \pi_k^*(W_k) \\
\gamma_k^* &= \gamma_k^*(W_k) \\
W_{k+1} &= W_k\left(1 + \pi_k^* \cdot (R_{k+1} - 1) - \gamma_k^*\right)
\end{aligned}$$

However, calculations simplify considerably in our setup, since the optimal γ_k^* and π_k^* do NOT depend on the current wealth v_k (see Remark below).

Remarks

1. Note that the dynamic programming principle only gives a *candidate* solution to (1). To confirm the optimality, one should check the sufficient condition of the Martingale Optimality Principle, i.e. show that $J(\vartheta^*, c^*)$ is a true martingale (where recall that $c_k^* = \gamma_k^* W_k^{v_0,\pi^*,\gamma^*}$ and $\vartheta_{k+1}^{*,l} = \frac{\pi_k^{*,l} W_k^{v_0,\pi^*,\gamma^*}}{X_k^l}$).

2. The (candidate) optimal proportions γ_k^*, π_k^* are deterministic. But there is still trading and stochastic consumption involved because

   $$c_k^* = \gamma_k^* W_k^{v_0,\pi^*,\gamma^*}$$

   and

   $$\vartheta_{k+1}^{*,l} = \frac{\pi_k^{*,l} W_k^{v_0,\pi^*,\gamma^*}}{X_k^l}$$

   still depend on the randomly evolving wealth and asset values.

Exercise 11.2 *The goal of the following exercises is to prove Bellman's Theorem for controlled dynamic systems. Questions 2 and 3 describe the basic setup needed for question 4. These results give more precise versions of those in the lecture for controlled dynamic systems and Markov Chains.*

Fix a probability space $(\Omega, \mathcal{F}, P)$ and consider a sequence of random variables $(X_n)_n$ taking values in a state space E endowed with a sigma-algebra $\mathcal{E}$. We suppose that the dynamics of X_n is given by

$$X_{n+1} = \phi_n(X_n, \epsilon_n) \tag{2}$$

where ϵ_n is a sequence of random variables valued in W (with sigma- algebra $\mathcal{W}$) with the property that ϵ_n is independent of $\sigma(X_0, \epsilon_1, \ldots, \epsilon_{n-1})$, and each $\phi_n : E \times W \mapsto E$ is a measurable function. For $x \in E$ and $A \in \mathcal{E}$ we define

$$P_n(x, A) = P\left(\phi_n(x, \epsilon_n) \in A\right) \tag{3}$$

Note that each P_n defines a transition kernel on $(E, \mathcal{E})$. Indeed,

- $\forall x \in E, A \mapsto P_n(x, A)$ is a probability measure on $(E, \mathcal{E})$.
- $\forall A \in \mathcal{E}, x \mapsto P_n(x, A)$ is measurable.

(a) Given a sequence (P_n) of transition kernels on $(E, \mathcal{E})$, a sequence of random variables (X_n) (valued in E) is called *inhomogeneous Markov Chain* with transition kernels (P_n), if $\forall A \in \mathcal{E}$ and $n \in \mathbb{N}$,

$$E\left[\mathbb{1}_{\{X_{n+1} \in A\}} | \sigma(X_0, \ldots, X_n)\right] = P_n(X_n, A) \tag{4}$$

Show that (X_n) defined by (2) defines an inhomogeneous Markov Chain with transition kernels given by (3).

We now introduce a very important notation: if P is a transition kernel on E and $f : E \mapsto \mathbb{R}_+$ is measurable, then we write for $x \in E$,

$$Pf(x) := \int_E f(y) P(x, dy) \tag{5}$$

Remark: the operator P defines a Feller semigroup. Feller semigroups are a very convinient way to talk about Markov Chains especially in a continous time setup.

(b) Show that if (X_n) is an inhomogeneous Markov Chain with transition kernels (P_n), and $f : E \mapsto \mathbb{R}_+$ is a measurable function, then $\forall n \in \mathbb{N}$,

$$E[f(X_{n+1}) | \sigma(X_0, \ldots, X_n)] = P_n f(X_n) \tag{6}$$

Solution 11.2

(a) First note by induction that X_n is $\sigma(X_0, \epsilon_0, ..., \epsilon_{n-1})$ measurable. Indeed it is clearly true for $n = 0$ and by induction, if X_n is $\sigma(X_0, \epsilon_0, ..., \epsilon_{n-1})$ measurable, then $X_n = f_n(X_0, \epsilon_0, ..., \epsilon_{n-1})$ for a certain measurable function f_n and so

$$X_{n+1} = \phi_n(X_n, \epsilon_n) = \phi_n\left(f_n(X_0, \epsilon_0, ..., \epsilon_{n-1}), \epsilon_n\right)$$

is $\sigma(X_0, \epsilon_0, \dots, \epsilon_n)$ measurable.

Since by assumption, ϵ_n is independent of $\sigma(X_0, \epsilon_0, \dots, \epsilon_{n-1})$, we get by the above result that ϵ_n is independent of $\sigma(X_0, \dots, X_n)$. Hence

$$\begin{aligned} E\left[\mathbb{1}_{\{X_{n+1}\in A\}}|\sigma(X_0, \dots, X_n)\right] &= E\left[\mathbb{1}_{\{\phi_n(X_n,\epsilon_n)\in A\}}|\sigma(X_0, \dots, X_n)\right] \\ &= P_n(X_n, A) \end{aligned}$$

where in the last equality we used the following proposition: if X, Y are random variables such that X is $\mathcal{G}$ measurable with respect to some sigma algebra $\mathcal{G}$ and Y is independent of $\mathcal{G}$, then

$$E[f(X,Y)|\mathcal{G}] = E[f(z,Y)]\Big|_{z=X}$$

(b) The equality holds for $f = \mathbb{1}_A$ for any $A \in \mathcal{E}$ by the previous question. By linearity it also holds for simple functions, i.e functions for the form $f = \sum_{k=1}^n a_k \mathbb{1}_{A_k}$. Finally it also holds for positive measurable functions since every such functions is the increasing limit of a sequence of simple functions.

Exercise 11.3 (Dynamic systems in discrete time: random controlled dynamic systems, i.e. Markov Decision Processes). We extend the setup of the previous question to allow for controls. We thus suppose that at each time n, the dynamics of the system (X_n) (taking values in $(E, \mathcal{E})$) can be controlled with a certain control U_n:

$$X_{n+1} = \phi_n(X_n, U_n, \epsilon_n) \tag{7}$$

where ϵ is again a sequence of independent random variables valued in W (with sigma-algebra $\mathcal{W}$) with the property that ϵ_n is independent of $\sigma(X_0, \epsilon_1, \dots, \epsilon_{n-1})$, and each ϕ_n is a measurable function. We also suppose that the random variable U_n representing the control at time n takes values in $(C, \mathcal{C})$ and that the process U is adapted to the natural filtration of X (i.e. for all $n \geq 0$, U_n is $\sigma(X_0, ..., X_n)$ measurable). This last assumption ensures that the control at time n only depends on the information available up to that time. Note that by the factorization lemma of measurable functions, adaptedness of U to the natural filtration of X is equivalent to the existence of a sequence of measurable maps $\eta_n : E^{n+1} \mapsto C$ such that

$$U_n = \eta_n(X_0, \dots, X_n)$$

We call the sequence $\eta = (\eta_n)_n$ a *strategy*. For $x \in E$, $u \in C$, and $A \in \mathcal{E}$, we define

$$P_n^{(u)}(x, A) = P\left(\phi_n(x, u, \epsilon_n) \in A\right) \tag{8}$$

Note that each $P_n^{(u)}$ defines a transition kernel on $(E, \mathcal{E})$. Indeed,

- $\forall x \in E, A \mapsto P_n^{(u)}(x, A)$ is a probability measure on $(E, \mathcal{E})$).
- $\forall A \in \mathcal{E}, x \mapsto P_n^{(u)}(x, A)$ is measurable.

(a) Show that (X_n) defined by (7) satsisfies $\forall A \in \mathcal{E}$ and $n \in \mathbb{N}$,

$$E\left[\mathbb{1}_{\{X_{n+1} \in A\}} | \sigma(X_0, \dots, X_n)\right] = P_n^{(U_n)}(X_n, A)$$

Note that we are not in a Markovian setup (as in the last exercise) since U_n can depend not only on X_n but also on the whole history of the process X up to time n.

We call (7) a random controlled dynamic system. The result of the previous sub-question tells us that the conditional law of X_{n+1} given $\sigma(X_0, ..., X_n)$ is $P_n^{(U_n)}(X_n, \cdot)$. Note that a random controlled dynamic system can thus be defined equivalently by specifying a strategy η and a family of transition kernels $(P_n^{(u)})$ such that $\forall A \in \mathcal{E}$, the map

$$(x, u) \mapsto P_n^{(u)}(x, A)$$

is measurable. Indeed, let $X_0 = x_0$ be fixed, and define recursively X_{n+1} by sampling from $P_n^{(U_n)}(X_n, \cdot)$ where $U_n = \eta_n(X_0, \dots, X_n)$. We will write $E_{x_0}^{\eta}$ for the expectation under the law of the process (X_n) starting at $X_0 = x_0$ and controlled by the strategy η. Recall from the last exercise, that we if P is a transition kernel on E and $f : E \mapsto \mathbb{R}_+$ is measurable, then we write for $x \in E$

$$Pf(x) := \int_E f(y)P(x, dy)$$

(b) Show that for all $f : E \mapsto \mathbb{R}_+$ positive measurable functions, we have $\forall n \in \mathbb{N}$,

$$E^{\eta}_{x_0}[f(X_{n+1})|\sigma(X_0 = x_0, \ldots, X_n)] = P_n^{(U_n)} f(X_n)$$

where $U_n = \eta_n(X_0, \ldots, X_n)$.

Solution 11.3

(a) The proof is very similar to the solution of the question 11.1.a). Again by induction one can see that X_n and U_n are $\sigma(X_0, \epsilon_0, ..., \epsilon_{n-1})$ measurable. Indeed it is again clearly true for $n = 0$ and by induction, if X_n and U_n are $\sigma(X_0, \epsilon_0, ..., \epsilon_{n-1})$ measurable, then there exists some measurable functions f_n and g_n such that $X_n = f_n(X_0, \epsilon_0, ..., \epsilon_{n-1})$ and $U_n = g_n(X_0, \epsilon_1, \ldots, \epsilon_{n-1})$

$$X_{n+1} = \phi_n(X_n, U_n, \epsilon_n) = \phi_n\left(f_n(X_0, \epsilon_0, ..., \epsilon_{n-1}), g_n(X_0, \epsilon_0, ..., \epsilon_{n-1}), \epsilon_n\right)$$

is $\sigma(X_0, \epsilon_0, \ldots, \epsilon_n)$ measurable and thus so is U_{n+1} since U is adapted to the natural filtration of X.

Since by assumption, ϵ_n is independent of $\sigma(X_0, \epsilon_0, \ldots, \epsilon_{n-1})$, we get by the above result that ϵ_n is independent of $\sigma(X_0, \ldots, X_n)$. Hence we have (using the same proposition as in question 11.1.a)),

$$\begin{aligned} E\left[\mathbb{1}_{\{X_{n+1} \in A\}} | \sigma(X_0, \ldots, X_n)\right] &= E\left[\mathbb{1}_{\{\phi_n(X_n, U_n, \epsilon_n) \in A\}} | \sigma(X_0, \ldots, X_n)\right] \\ &= P_n^{(U_n)}(X_n, A) \end{aligned}$$

(b) The solution follows the same ideas as question 11.1.b). The equality holds for $f = \mathbb{1}_A$ for any $A \in \mathcal{E}$ by the previous question. By linearity it also holds for simple functions, i.e functions for the form $f = \sum_{k=1}^n a_k \mathbb{1}_{A_k}$. Finally it also holds for positive measurable function since every such functions is the increasing limit of a sequence of simple functions.

Exercise 11.4 (Dynamic Programming: Bellman's Theorem) Consider a random controlled dynamic system starting at $X_0 = x_0$ (see previous exercise). Fix a time horizon N, intermediate cost functions $c_k : E \times C \mapsto \mathbb{R} \cup \{+\infty\}$ for $k = 0, \dots, N-1$ and a terminal cost function $\gamma : E \mapsto \mathbb{R} \cup \{+\infty\}$. In this exercise we are interested in minimizing the total expected cost

$$E\left[\sum_{k=0}^{N-1} c_k(X_k, U_k) + \gamma(X_N)\right] \tag{9}$$

over the sequence of controls $U_0, \dots, U_{N-1}$.
We break down this (unconditional) optimization problem into a series of much simpler conditional optimization problems. Define by backwards induction

$$\begin{aligned} O_N(x) &= \gamma(x) \qquad \forall x \in E \\ O_n(x) &= \min_{u \in C} \Big(c_n(x,u) + (P_n^{(u)} O_{n+1})(x) \Big) \qquad \forall x \in E \qquad n = N-1, \dots, 1, 0 \end{aligned} \tag{10}$$

where recall that we have showed in the last exercise that

$$(P_n^{(U)} O_{n+1})(X_n) = E_{x_0}^{\eta}\left[O_{n+1}(X_{n+1}) | \sigma(X_0, \dots, X_n)\right]$$

where η is the strategy corresponding to the controls U.
The goal of this exercise is to show Bellman's Theorem

Theorem 1. *Suppose that for all $x \in E$ and $0 \le n \le N$, the optimization problem* (10) *admits a solution and write $u_n^{min}(x)$ for the mimizer. Moreover suppose that the map $u_n^{min} : E \mapsto C$ is measurable. Then there exists a strategy η which minimizes* (9) *and this strategy is given by $\eta^{min} := (u_n^{min})_{n=0,\dots,N-1}$. Moreover, for all n, we have*

$$O_n(X_n) = \min_{\textit{strategy } \eta} E_{x_0}^{\eta}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N) | \sigma(X_0, \dots, X_n)\right]$$

In particular,

$$O_0(x_0) = \min_{\textit{strategy } \eta} E_{x_0}^{\eta}\left[\sum_{k=0}^{N-1} c_k(X_k, U_k) + \gamma(X_N)\right]$$

is the optimal value of (9).

To prove this theorem you are asked to show that at all time steps n we have:

(a)

$$O_n(X_n) = E_{x_0}^{\eta^{min}}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N) | \sigma(X_0, \dots, X_n)\right]$$

where recall that $\eta^{min} := (u_n^{min})_{n=0,\dots,N-1}$.

(b) for all strategy η,

$$E_{x_0}^{\eta^{min}}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N)|\sigma(X_0, \ldots, X_n)\right] \leq E_{x_0}^{\eta}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N)|\sigma(X_0, \ldots, X_n)\right]$$

Hint: Use backwards induction for both questions

Solution 11.4 As suggested by the hint, we use bakward induction for both questions.

(a) It is clearly true for $n = N$:

$$O_N(X_N) = \gamma(X_N) = E_{x_0}^{\eta^{min}}\left[\gamma(X_N)|\sigma(X_0, \ldots, X_N)\right]$$

Now we suppose that the desired property holds for $n+1$ and show that it still holds for n

$$\begin{aligned}
O_n(X_n) &= c_n(X_n, U_n^{(min)}(X_n)) + (P_n^{(U_n^{(min)}(X_n))} O_{n+1})(X_n) \quad \textit{(by definition)}\\
&= c_n(X_n, U_n^{(min)}(X_n)) + E_{x_0}^{\eta^{min}}\left[O_{n+1}(X_{n+1})|\sigma(X_0, \ldots, X_n)\right]\\
&\quad \textit{(using the result from question 2)}\\
&= c_n(X_n, U_n^{(min)}(X_n))\\
&\quad + E_{x_0}^{\eta^{min}}\left[E_{x_0}^{\eta^{min}}\left[\sum_{k=n+1}^{N-1} c_k(X_k, U_k) + \gamma(X_N)|\sigma(X_0, \ldots, X_{n+1})\right]|\sigma(X_0, \ldots, X_n)\right]\\
&\quad \textit{(by induction hypotheses)}\\
&= c_n(X_n, U_n^{(min)}(X_n))\\
&\quad + E_{x_0}^{\eta^{min}}\left[\sum_{k=n+1}^{N-1} c_k(X_k, U_k) + \gamma(X_N)|\sigma(X_0, \ldots, X_n)\right]\\
&\quad \textit{(by tower property)}\\
&= E_{x_0}^{\eta^{min}}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N)|\sigma(X_0, \ldots, X_n)\right]
\end{aligned}$$

(b) Fix a strategy η corresponding to controls U. The claimed result is clearly true for $n = N$:

$$O_N(X_N) = \gamma(X_N) = E_{x_0}^{\eta}\left[\gamma(X_N)|\sigma(X_0, \ldots, X_N)\right]$$

Now we suppose that the desired property holds for $n+1$ and show that it still holds for n

$$
\begin{aligned}
O_n(X_n) &= c_n(X_n, U_n^{(min)}(X_n)) + (P_n^{(U_n^{(min)}(X_n))} O_{n+1})(X_n) \quad \textit{(by definition)} \\
&\leq c_n(X_n, U_n(X_n)) + (P_n^{(U_n(X_n))} O_{n+1})(X_n) \quad \textit{(by definition of optimal strategy)} \\
&= c_n(X_n, U_n(X_n)) + E_{x_0}^{\eta}\left[O_{n+1}(X_{n+1}) | \sigma(X_0, \dots, X_n)\right] \\
&\quad \textit{(using the result from question 3)} \\
&\leq c_n(X_n, U_n(X_n)) \\
&\qquad + E_{x_0}^{\eta}\left[E_{x_0}^{\eta}\left[\sum_{k=n+1}^{N-1} c_k(X_k, U_k) + \gamma(X_N) | \sigma(X_0, \dots, X_{n+1})\right] | \sigma(X_0, \dots, X_n)\right] \\
&\qquad \textit{(by induction hypotheses + result from question (a))} \\
&= c_n(X_n, U_n(X_n)) \\
&\qquad + E_{x_0}^{\eta}\left[\sum_{k=n+1}^{N-1} c_k(X_k, U_k) + \gamma(X_N) | \sigma(X_0, \dots, X_n)\right] \\
&\qquad \textit{(by tower property)} \\
&= E_{x_0}^{\eta}\left[\sum_{k=n}^{N-1} c_k(X_k, U_k) + \gamma(X_N) | \sigma(X_0, \dots, X_n)\right]
\end{aligned}
$$