Portfolios

A Markov chain of the following form is called a \textit{random walk with reflecting barriers}.

This is a rough model of a volatile stock which ranges randomly between some upper bound at which it is judged overpriced and listed as a “sell” and lower bound at which it is judged undervalued and listed as a “buy”.

A stock portfolio consists of two investments a stock and a bond. The \textit{stock} is a volatile stock which is a random walk with reflecting barriers as above with probability $p$ of a unit increase, probability $1-p$ of decrease except when at a barrier. The \textit{bond} is a government bond which consistently earns an interest of $\beta$ and hence grows each year by a factor of $1+\beta$. 
At any time, you can withdraw funds from either investment.

You can either keep the withdrawal or deposit it in the other investment.

Your goal is to maximize the total amount withdrawn.

The finite horizon case has a fixed terminal time at which time all funds are withdrawn from both investments.

In the discounted infinite horizon case, you maximize your total withdrawals over an infinite amount of time with some discount rate $\alpha$. $\alpha < 1$. 
The state at any time is a vector \((x, k, y, l)\) where
\(x\) = the current price for one share of the stock.
\(k\) = the number of shares you own of this volatile stock.
\(y\) = the current price for one share of the bond.
\(l\) = the number of shares you own of the bond.

The *value* of the stocks and bonds is ??

In the finite horizon case, the final reward is the value of the stocks and bonds at time \(T = R(x, k, y, l) = ??\)

An *action* is a vector \((a, b)\) where \(a\) is the number of stocks sold (or equivalently, withdrawn) and \(b\) is the number of bond shares sold.

The reward \(r(x, k, y, l, a, b) = ?? = the current withdrawal = the current sale.\)
The state at any time is a vector \( (x, k, y, l) \) where
\[
x = \text{the current price for one share of the stock.}
\]
\[
k = \text{the number of shares you own of this volatile stock.}
\]
\[
y = \text{the current price for one share of the bond.}
\]
\[
l = \text{the number of shares you own of the bond.}
\]

The \textit{value} of the stocks and bonds is \( kx + ly \). In the finite horizon case, the final reward is the value
\[
R(x, k, y, l) = kx + ly
\]
of the stocks and bonds at time \( T \).

An \textit{action} is a vector \( (a, b) \) where \( a \) is the number of stocks sold (or equivalently, withdrawn) and \( b \) is the number of bond shares sold.

The reward \( r(x, k, y, l, a, b) = ax + by = \text{the current withdrawal} \) = the current sale.
Buying is the negative of selling. So \( a = 3 \) means three stock shares are sold and \( b = -2 \) means two bond shares are bought, i.e., a negative withdrawal is a buy.

A policy is a pair of functions \((u^a(x, k, y, l), u^b(x, k, y, l))\) where \( u^a(x, k, y, l) = \text{the number of stock shares to sell} \) and \( u^b(x, k, y, l) = \text{the number of bond shares to sell} \). Each action depends on the current state \((x, k, y, l)\). Nonstationary policies (used in the finite horizon case) also depend on the current time \(n\): \((u^n_a(x, k, y, l), u^n_b(x, k, y, l))\).

If you the amount you buy exceeds the amount you sell, then \(ax + by < 0\). We now assume this doesn’t happen.

**Assumption.** No new money is used to purchase stocks or bonds. When you buy one investment (stock or bond) the money used for the purchase is withdrawn from the other investment.

Thus the total sale \(ax + by\) cannot be negative. Also, the current sale cannot exceed the current value. Thus \(0 \leq ax + by \leq kx + ly\).

Should you ever sell some but not all shares?
Now calculate the transition probability $T_{(a,b)}( (x, k, y, l), (x^*, k^*, y^*, l^*))$ for going from state $(x, k, y, l)$ to state $(x^*, k^*, y^*, l^*)$ if the action you take is $(a, b)$. With probability one we must have:

- $k^* = ??,$
- $l^* = ??,$
- $y^* = ??.$

Recall:
- $x =$ the current price for one share of the stock.
- $k =$ the number of shares he owns of this volatile stock.
- $y =$ the current price for one share of the bond.
- $l =$ the number of shares he owns of the bond.
- $a =$ the number of stocks sold.
- $b =$ the number of bond shares sold.
- A bond grows each year by a factor of $1+\beta$. 
Now calculate the transition probability
\[ T_{(a,b)}((x, k, y, l), (x^*, k^*, y^*, l^*)) \] for going
from state \((x, k, y, l)\)
to state \((x^*, k^*, y^*, l^*)\) if the action you take is \((a, b)\).
With probability one we must have:
\[ k^* = k - a, \quad l^* = l - b, \quad y^* = (1 + \beta)y. \]
If the three conditions are true, then the transition probability
\[ T_{(a,b)}((x, k, y, l), (x^*, k^*, y^*, l^*)) = \] the transition probability
\[ T(x, x^*) \] that the price of the volatile stock will go from \(x\) to \(x^*\). This is the probability given by the random walk.

In the barrier cases,
\[ x = 0 \rightarrow x^* = 1 \text{ and} \]
\[ x = N \rightarrow x^* = N - 1 \text{ have probability 1.} \]
Otherwise
\[ x^* = x + 1 \text{ has probability } p \text{ and} \]
\[ x^* = x - 1 \text{ has probability } 1 - p. \]
Suppose $X_0, K_0, Y_0, L_0; X_1, K_1, Y_1, L_1; \ldots X_2, K_2, Y_2, L_2; \ldots$ is the sequence of random variables for the states at times $n = 0, 1, 2, \ldots$, where $X_n, K_n, Y_n, L_n$ are the values of $x, k, y, l$ at time $n$ when a policy $u$ is followed. Let $U_n^a = u^a(X_n, K_n, Y_n, L_n)$, $U_n^b = u^b(X_n, K_n, Y_n, L_n)$ be the actions taken at time $n$.

**Finite horizon portfolios:**
The value function for the finite horizon problem with final time $T$ and policy $u$ and initial state $X_{init} = (x, k, y, l)$ is:

$$V(x, k, y, l, u) = E[R(X_T, K_T, Y_T, L_T) + \sum_{n=0}^{T-1} r(X_n, K_n, Y_n, L_n, U_n^a, U_n^b) | X_0 = X_{init}]$$

$$= E[?? + \sum_{n=0}^{T-1} (U_n^a X_n + U_n^b Y_n) | X_0 = X_{init}].$$

If $U_{n-1}^a = a$ and $U_{n-1}^b = b$, then the recursion equation is

$$V_{n-1}(x, k, y, l, u)$$

$$= ?? + \sum T_{(a,b)}(x, x^*)( ?? )$$

The optimal value $V(x, k, y, l) = \max_u V(x, k, l, u)$ has the dynamic programming equation

$$V_{n-1}(x, k, y, l)$$

$$= \max_{a,b} \{ ?? + \sum_{x*} T(x, x^*)( ?? ) \}$$
Suppose $X_0, K_0, Y_0, L_0; X_1, K_1, Y_1, L_1; X_2, K_2, Y_2, L_2; \ldots$ is the sequence of random variables for the states at times $n = 0, 1, 2, \ldots$, where $X_n, K_n, Y_n, L_n$ are the values of $x, k, y, l$ at time $n$ when a policy $u$ is followed. Let $U_n^a = u^a(X_n, K_n, Y_n, L_n)$, $U_n^b = u^b(X_n, K_n, Y_n, L_n)$ be the actions taken at time $n$.

**FINITE HORIZON PORTFOLIOS:**
The value function for the finite horizon problem with final time $T$ and policy $u$ and initial state $X_{init} = (x, k, y, l)$ is:

\[
V(x, k, y, l, u) = E[R(X_T, K_T, Y_T, L_T) + \sum_{n=0}^{T-1} r(X_n, K_n, Y_n, L_n, U_n^a, U_n^b) | X_0 = X_{init}] \\
= E[K_TX_T + L_TY_T + \sum_{n=0}^{T-1} (U_n^aX_n + U_n^bY_n) | X_0 = X_{init}].
\]

If $U_{n-1}^a = a$ and $U_{n-1}^b = b$, then the recursion equation is

\[
V_{n-1}(x, k, y, l, u) = ax + by + \sum_{x^*} T(x, x^*)V_n(x^*, k - a, (1 + \beta)y, l - b, u).
\]

The optimal value $V(x, k, y, l) = \max_u V(x, k, l, u)$ has the dynamic programming equation

\[
V_{n-1}(x, k, y, l) = \max_{a, b} \{ ax + by + \sum_{x^*} T(x, x^*)V_n(x^*, k - a, (1 + \beta)y, l - b) \}.
\]
Should you ever pocket money from a sale as opposed to transferring it to the other stock?

**INFINITE HORIZON PORTFOLIOS:**

The value function with discount $\alpha$, policy $u$ and initial state $X_{init} = (x, k, y, l)$ is:

$$W(x, k, y, l, u) = E[\sum_{n=0}^{\infty} \alpha^n r(X_n, K_n, Y_n, L_n, U^a_n, U^b_n) | X_0 = X_{init}] = E[\sum_{n=0}^{\infty} \alpha^n (U^a_n X_n + U^b_n Y_n) | X_0 = X_{init}].$$

For $U^a_0 = a, U^b_0 = b$ the recursion equation is

$$W(x, k, y, l, u) = ?? + \sum_{x^*} T(x, x^*) (??)$$

The optimal value’s dynamic programming equation is

$$W(x, k, y, l) = \max_{a, b} \{ ?? + \sum_{x^*} T(x, x^*) (??) \}.$$
Should you ever pocket money from a sale as opposed to transferring it to the other stock?

**Infinite horizon portfolios:**

The value function with discount $\alpha$, policy $u$ and initial state $X_{init} = (x, k, y, l)$ is: $W(x, k, y, l, u) = E[\sum_{n=0}^{\infty} \alpha^n r(X_n, K_n, Y_n, L_n, U^a_n, U^b_n) | X_0 = X_{init}]$

$= E[\sum_{n=0}^{\infty} \alpha^n (U^a_n X_n + U^b_n Y_n) | X_0 = X_{init}]$.

For $U^a_0 = a, U^b_0 = b$ the recursion equation is

$W(x, k, y, l, u) = ax + by + \sum_{x^*} T(x, x^*) W(x^*, k - a, (1 + \beta)y, l - b, u)$.

The optimal value’s dynamic programming equation is

$W(x, k, y, l) = \max_{a, b} \{ ax + by + \sum_{x^*} T(x, x^*) W(x^*, k - a, (1 + \beta)y, l - b) \}$. 
Portfolios with no buying, finite horizon

In our model, the stock price $x$ is independent of the number of shares $k$ one has or the number $a$ one sells (for large shareholders, this is not true, selling a substantial fraction of a company’s stock usually depresses the price).

Now add a further restriction that neither stocks nor bonds may be bought; the only possible action is to sell shares of the stock or bond. In the above case, the stocks and bonds were dependent since buying shares of one required selling shares of the other. Now the two are independent. Hence we can find the optimal actions for each separately. First consider the finite horizon case with a fixed terminal time $T$ when all shares are sold.
**Bond Case.** The bonds earn an interest of $\beta$. We have $l$ bonds each valued at $y$ dollars.
   
   If we sell $a$ shares now, we get $??$
   
   If we sell them $n$ years from now, we get $??$
   
   Should we wait until $x = N$ before selling?

**Stock Case.** We have $k$ bonds worth $x$ dollars each.
   
   Selling $a$ shares at price $x$ gives a reward of $ax$.
   
   If there are $k$ shares at time $T$ each with price $x$, the final reward is $??$.
   
   The possible actions are selling $a \in A = \{0, 1, 2, ..., k\}$ shares.
   
   A policy $u$ for the stock determines the number of shares $a = u(x, k)$ to sell.
**Bond Case.** The bonds earn an interest of $\beta$. We have $l$ bonds each valued at $y$ dollars.

If we sell $a$ shares now, we get $ay$.

If we sell them $n$ years from now, we get $(1 + \beta)^n ay$.

The optimal policy is to sell no bond shares before the final time $T$ at which time we sell all $l$ shares and get $(1 + \beta)^T ly$.

**Stock Case.** We have $k$ bonds worth $x$ dollars each.

Selling $a$ shares at price $x$ gives a reward of $ax$.

If there are $k$ shares at time $T$ each with price $x$, the final reward is $kx$.

The possible actions are selling $a \in A = \{0, 1, 2, \ldots, k\}$ shares.

A policy $u$ for the stock determines the number of shares $a = u(x, k)$ to sell.

Should you ever sell some but not all shares? Why not wait until $x = N$ to sell the stocks?
Let $X_0, K_0, X_1, K_1, X_2, K_2, \ldots$ be a sequence of states for times $n=0, 1, 2, \ldots$. Let $U_0, U_1, U_2, \ldots$ be the sequence of stock-selling actions $U_i = u(X_i, K_i)$ determined by the policy $u$. Then the policy’s finite horizon value function with an initial state $(x, k)$ is:

$$V(x, k, u) = E[R(X_T, K_T) + \sum_{n=0}^{T-1} r(X_n, K_n, U_n) \mid X_0 = (x, k)]$$

$$= E[K_TX_T + \sum_{n=0}^{T-1} (U_nX_n) \mid X_0 = (x, k)]$$

The optimal value function is $V(x, k) = \max_u V(x, k, u)$. Calculate the optimal value function $V_n(x, k)$ and nonstationary optimal policy $u_n(x, k)$ at time $n$ by working backward from the known time $T$ (Lecture 26).

$$V_{n-1}(x, k)$$

$$= \max_a [r(x, k, a) + \sum_{x^*, k^*} T_a(x, k, x^*, k^*)V_n(x^*, k^*)]$$

$$= \max_a [a x + \sum_{x^*} T(x, x^*)V_n(x^*, k - a)]$$

The transition probabilities $T$ for the price $x$ of the volatile stock are pictured in the Markov chain above. It is independent of the number $a$ of shares the action sells.
**Dynamic programming equation.**

\[ V_T(x, k) = R(x, k) = kx \]

\[ V_{n-1}(x, k) = \max_a [ax + \sum x^* T(x, x^*) V_n(x^*, k - a)] \]

where the max is taken over actions \( a = 0, 1, \ldots, k \).

There are three cases depending on whether \( x \) is

- the first state 0 of the Markov chain,
- a middle state in \([1, N-1]\), or
- the last state \( N \).

Simplify the dynamic recursion equation:

\[ V_{n-1}(x, k) = \max_a [ax + \sum x^* T(x, x^*) V_n(x^*, k - a)] \]

Find any obvious actions \( a \).

**Case** \( x = 0 \).

\[ V_{n-1}(0, k) = ?? \]

**Case** \( x \in [1, N-1] \).

\[ V_{n-1}(x, k) = ?? \]

**Case** \( x = N \).

\[ V_{n-1}(N, k) = ?? \]
There are three cases depending on whether \( x \) is
the first state 0 of the Markov chain,
a middle state in \([1, N-1]\), or
the last state \( N \).

Simplify the dynamic recursion equation:
\[
V_{n-1}(x, k) = \max_a [ax + \sum_{x^*} T(x, x^*) V_n(x^*, k - a)].
\]
Find any obvious actions \( a \).

**Case \( x = 0 \).**  \( V_{n-1}(0, k) = \) (What is the obvious action \( a = ? \))
\[
= \max_a [a0 + T(0, 1) V_n(1, k - a)]
\]
\[
= \max_a [V_n(1, k - a)] = V_n(1, k) \quad \text{(Why?)}
\]

**Case \( x \in [1, N-1] \).**  \( V_{n-1}(x, k) = \)
\[
= \max_a [ax + \sum_{x^*} T(x, x^*) V_n(x^*, k - a)]
\]
\[
= \max_a [ax + pV_n(x + 1, k - a) + (1 - p)V_n(x - 1, k - a)]
\]

**Case \( x = N \).**  \( V_{n-1}(N, k) = \) (What is the obvious action \( a = ? \))
\[
= \max_a [aN + T(N, N - 1) V_n(N - 1, k - a)]
\]
\[
= \max_a [aN + V_n(N - 1, k - a)] = kN \quad \text{(Why?)}
\]