

From (13.29), we can formally write

$$(dX_t)^2 = f^2(dt)^2 + G^2(dz_t)^2 + 2fGdz_tdt, \quad (13.32)$$

$$dX_tdt = f(dt)^2 + Gdz_tdt. \quad (13.33)$$

The exact meaning of these expressions comes from the theory of stochastic calculus; see Arnold (1974, Chapter 5), Durrett (1996) or Karatzas and Shreve (1997). For our purposes, it is sufficient to know the multiplication rules of the stochastic calculus:

$$(dz_t)^2 = dt, \quad dz_t \cdot dt = 0, \quad dt^2 = 0. \quad (13.34)$$

Substitute (13.31) into (13.30) and use (13.32), (13.33), (13.34), and the property that $E[z_t] = 0$ to obtain

$$V = \max_u E \left[Fdt + V + V_tdt + V_xfdt + \frac{1}{2}V_{xx}G^2dt + o(dt) \right]. \quad (13.35)$$

Note that we have suppressed the arguments of the functions involved in (13.35).

Cancelling the the term V on both sides of (13.35), dividing the remainder by dt , and letting $dt \rightarrow 0$, we obtain the Hamilton-Jacobi-Bellman equation

$$0 = \max_u [F + V_t + V_xf + \frac{1}{2}V_{xx}G^2] \quad (13.36)$$

for the value function $V(t, x)$ with the boundary condition

$$V(x, T) = S(x, T). \quad (13.37)$$

In the next section we shall apply this theory of stochastic optimal control to a simple stochastic production inventory problem treated by Sethi and Thompson (1981).

13.3 A Stochastic Production Planning Model

Consider a factory producing a homogeneous good and having an inventory warehouse. Define the following quantities:

- X_t = the inventory level at time t (state variable),
- U_t = the production rate at time t (control variable),

(13.40), we restate (13.39) as

$$\max E \left\{ \int_0^T -(U_t^2 + X_t^2) dt + BX_T \right\}. \quad (13.41)$$

Let $V(x, t)$ denote the expected value of the objective function from time t to the horizon T with $X_t = x$ and using the optimal policy from t to T . The function $V(x, t)$ is referred to as the value function, and it satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \max_u [-(u^2 + x^2) + V_t + V_x(u - S) + \frac{1}{2}\sigma^2 V_{xx}] \quad (13.42)$$

with the boundary condition

$$V(x, T) = Bx. \quad (13.43)$$

Note that these are applications of (13.36) and (13.37) to the production planning problem.

It is now possible to maximize the expression inside the bracket of (13.28) with respect to u by taking its derivative with respect to u and setting it to zero. This procedure yields

$$u(x, t) = \frac{V_x(x, t)}{2}. \quad (13.44)$$

Substituting (13.44) into (13.42) yields the equation

$$0 = \frac{V_x^2}{4} - x^2 + V_t - SV_x + \frac{1}{2}\sigma^2 V_{xx}, \quad (13.45)$$

known as the Hamilton-Jacobi equation. This is a partial differential equation which must be satisfied by the value function $V(x, t)$ with the boundary condition (13.43). The solution of (13.45) is considered in the next section.

Remark 13.1 It is important to remark that if production rate were restricted to be nonnegative, then (13.44) would be changed to

$$u(x, t) = \max \left[0, \frac{V_x(x, t)}{2} \right]. \quad (13.46)$$

Substituting (13.46) into (13.43) would give us a partial differential equation which must be solved numerically. We shall not consider (13.46) further in this chapter.

13.3.1 Solution for the Production Planning Problem

To solve equation (13.45) we let

$$V(x, t) = Q(t)x^2 + R(t)x + M(t). \quad (13.47)$$

Then,

$$V_t = \dot{Q}x^2 + \dot{R}x + \dot{M}, \quad (13.48)$$

$$V_x = 2Qx + R, \quad (13.49)$$

$$V_{xx} = 2Q, \quad (13.50)$$

where \dot{Y} denotes dY/dt . Substituting (13.48) in (13.45) and collecting terms gives

$$x^2[\dot{Q} + Q^2 - 1] + x[\dot{R} + RQ - 2SQ] + \dot{M} + \frac{R^2}{2} - RS + \sigma^2Q = 0. \quad (13.51)$$

Since (13.51) must hold for any value of x , we must have

$$\dot{Q} = 1 - Q^2, \quad Q(T) = 0, \quad (13.52)$$

$$\dot{R} = 2SQ - RQ, \quad R(T) = B, \quad (13.53)$$

$$\dot{M} = RS - \frac{R^2}{4} - \sigma^2Q, \quad M(T) = 0, \quad (13.54)$$

where the boundary conditions for the system of simultaneous differential equations (13.52), (13.53), and (13.54) are obtained by comparing (13.47) with the boundary condition $V(x, T) = Bx$ of (13.43).

To solve (13.52), we expand $\dot{Q}/(1 - Q^2)$ by partial fractions to obtain

$$\frac{\dot{Q}}{2} \left[\frac{1}{1 - Q} + \frac{1}{1 + Q} \right] = 1,$$

which can be easily integrated. The answer is

$$Q = \frac{y - 1}{y + 1}, \quad (13.55)$$

where

$$y = e^{2(t-T)}. \quad (13.56)$$

Since S is assumed to be a constant, we can reduce (13.53) to

$$\dot{R}^0 + R^0Q = 0, \quad R^0(T) = B - 2S$$

by the change of variable defined by $R^0 = R - 2S$. Clearly the solution is given by

$$\log R^0(T) - \log R^0(t) = - \int_t^T Q(\tau) d\tau,$$

which can be simplified further to obtain

$$R = 2S + \frac{2(B - 2S)\sqrt{y}}{y + 1}. \quad (13.57)$$

Having obtained solutions for R and Q , we can easily express (13.54) as

$$M(t) = - \int_t^T [R(\tau)S - (R(\tau))^2/4 - \sigma^2 Q(\tau)] d\tau. \quad (13.58)$$

The optimal control is defined by (13.44), and the use of (13.55) and (13.57) yields

$$u^* = V_x/2 = Qx + R/2 = S + \frac{(y - 1)x + (B - 2S)\sqrt{y}}{y + 1}. \quad (13.59)$$

Remark 13.2 The optimal production rate in (13.59) equals the demand rate plus a correction term which depends on the level of inventory and the distance from the horizon time T . Since $(y - 1) < 0$ for $t < T$, it is clear that for lower values of x , the optimal production rate is likely to be positive. However, if x is very high, the correction term will become smaller than $-S$, and the optimal control will be negative. In other words, if inventory level is too high, the factory can save money by disposing a part of the inventory resulting in lower holding costs.

Remark 13.3 If the demand rate S were time-dependent, it would have changed the solution of (13.53). Having computed this new solution in place of (13.57), we can once again obtain the optimal control as $u^* = Qx + R/2$.

Remark 13.4 Note that when $T \rightarrow \infty$, we have $y \rightarrow 0$ and

$$u^* \rightarrow S - x, \quad (13.60)$$

but the undiscounted objective function value (13.41) in this case becomes $-\infty$. Clearly, any other policy will render the objective function value to be $-\infty$. In a sense, the optimal control problem becomes ill-posed. One way to get out of this difficulty is to impose a nonzero discount rate. This is carried out in Sethi and Thompson (1980).

time t . Thus, the term in the square bracket represents the total discounted profits on a sample path. The objective in (13.61) is, therefore, to maximize the expected value of the total discounted profits.

This model is a modification as well as a stochastic extension of the optimal control formulation of the Vidale-Wolfe advertising model presented in (7.39). The Itô equation in (13.61) modifies the Vidale-Wolfe dynamics (7.22) by replacing the term $ru(1-x)$ by $rU_t\sqrt{1-X_t}$ and adding a diffusion term $\sigma(X_t)dz_t$ on the right-hand side. Furthermore, we replace the linear cost of advertising u in (7.39) by a quadratic cost of advertising U_t^2 in (13.61). We also relax the control constraint $0 \leq u \leq Q$ in (7.39) to simplify $U_t \geq 0$. The addition of the diffusion term yields a stochastic optimal control problem as expressed in (13.61).

An important consideration in choosing the function $\sigma(x)$ should be that the solution X_t to the Itô equation in (13.61) remains inside the interval $[0, 1]$. Merely requiring that the initial condition $x_0 \in [0, 1]$, as in Section 7.2.1, is no longer sufficient in the stochastic case. Additional conditions need to be imposed. It is possible to specify these conditions by using the theory presented by Gihman and Skorohod (1972) for stochastic differential equations on a finite spatial interval. In our case, the conditions boil down to the following, in addition to $x_0 \in (0, 1)$, which has been assumed already in (13.61):

$$\sigma(x) > 0, \quad x \in (0, 1) \text{ and } \sigma(0) = \sigma(1) = 0. \quad (13.62)$$

It is possible to show that for any feedback control $u(x)$ satisfying

$$u(x) \geq 0, \quad x \in (0, 1], \text{ and } u(0) > 0, \quad (13.63)$$

the Itô equation in (13.61) will have a solution X_t such that $0 < X_t < 1$, *almost surely* (i.e., with probability 1). Since our solution for the optimal advertising $u^*(x)$ would turn out to satisfy (13.63), we will have the optimal market share X_t^* lie in the interval $(0, 1)$.

Let $V(x)$ denote the value function for the problem, i.e., $V(x)$ is the expected value of the discounted profits from time t to infinity. When $X_t = x$ and an optimal policy U_t^* is followed from time t onwards. Note that since $T = \infty$, the future looks the same from any time t , and therefore the value function does not depend on t . It is for this reason we have defined the value function as $V(x)$, rather than $V(x, t)$ as in the previous section.

Using now the principle of optimality as in Section 13.2, we can write the HJB equation as

$$\rho V(x) = \max_u \left[\pi x - u^2 + V_x(ru\sqrt{1-x} - \delta x) + V_{xx}\sigma^2(x)/2 \right]. \quad (13.64)$$

Maximization of the RHS of (13.64) can be accomplished by taking its derivative with respect to u and setting it to zero. This gives

$$u(x) = \frac{rV_x\sqrt{1-x}}{2}. \quad (13.65)$$

Substituting of (13.65) in (13.64) and simplifying the resulting expression yields the HJB equation

$$\rho V(x) = \pi x + \frac{V_x^2 r^2 (1-x)}{4} - V_x \delta x + \frac{1}{2} \sigma^2(x) V_{xx}. \quad (13.66)$$

As shown in Sethi (1983b), a solution of (13.66) is

$$V(x) = \bar{\lambda}x + \frac{\bar{\lambda}^2 r^2}{4\rho}, \quad (13.67)$$

where

$$\bar{\lambda} = \frac{\sqrt{(\rho + \delta)^2 + r^2 \pi} - (\rho + \delta)}{r^2/2}, \quad (13.68)$$

as derived in Exercise 7.40. In Exercise 13.4, you are asked verify that (13.67) and (13.68) solve the HJB equation (13.66).

We can now obtain the explicit formula for the optimal feedback control as

$$u^*(x) = \frac{r\bar{\lambda}\sqrt{1-x}}{2}. \quad (13.69)$$

Note that $u^*(x)$ satisfies the conditions in (13.63).

As in Exercise 7.40, it is easy to characterize (13.69) as

$$U_t^* = u^*(X_t) = \begin{cases} > \bar{u} & \text{if } X_t < \bar{x}, \\ = \bar{u} & \text{if } X_t = \bar{x}, \\ < \bar{u} & \text{if } X_t > \bar{x}, \end{cases} \quad (13.70)$$

where

$$\bar{x} = \frac{r^2 \bar{\lambda} / 2}{r^2 \bar{\lambda} / 2 + \delta} \quad (13.71)$$

and

$$\bar{u} = \frac{r\bar{\lambda}\sqrt{1-\bar{x}}}{2}, \quad (13.72)$$

as given in (7.48).

modelled by an Itô equation, namely,

$$\frac{dP_t}{P_t} = \alpha dt + \sigma dz_t, \quad P_0 \text{ given}, \quad (13.74)$$

or simply,

$$dP_t = \alpha P_t dt + \sigma P_t dz_t, \quad P_0 \text{ given}, \quad (13.75)$$

where α is the average rate of return on stock, σ is the standard deviation associated with the return, and z_t is a standard Wiener process.

Remark 13.6 The LHS in (13.74) can be written also as $d \ln P_t$. Another name for the process z_t is *Brownian Motion*. Because of these, the price process P_t given by (13.74) is often referred to as a *logarithmic Brownian Motion*.

In order to complete the formulation of Rich's stochastic optimal control problem, we need the following additional notation:

- W_t = the wealth at time t ,
- C_t = the consumption rate at time t ,
- Q_t = the fraction of the wealth invested in stock at time t ,
- $1 - Q_t$ = the fraction of the wealth kept in the savings account at time t ,
- $U(c)$ = the utility of consumption when consumption is at the rate c ; the function $U(c)$ is assumed to be increasing and concave,
- ρ = the rate of discount applied to consumption utility,
- B = the bankruptcy parameter to be explained later.

Next we develop the dynamics of the wealth process. Since the investment decision Q is unconstrained, it means Rich is allowed to buy stock as well as to sell it short. Moreover, Rich can deposit in, as well as borrow money from, the savings account at the rate r .

While it is possible to obtain rigorously the equation for the wealth process involving an intermediate variable, namely, the number N_t of shares of stock owned at time t , we shall not do so. Instead, we shall write the wealth equation informally as

$$\begin{aligned} dW_t &= Q_t W_t \alpha dt + Q_t W_t \sigma dz_t + (1 - Q_t) r W_t dt - C_t dt \\ &= (\alpha - r) Q_t W_t dt + (r W_t - C_t) dt + \sigma Q_t W_t dz_t, \quad W_0 \text{ given}, \end{aligned} \quad (13.76)$$

i.e.,

$$q(x) = -\frac{(\alpha - r)V_x}{x\sigma^2 V_{xx}}, \quad (13.82)$$

and

$$c(x) = \frac{1}{V_x}. \quad (13.83)$$

Substituting (13.82) and (13.83) in (13.79) allows us to remove the max operator from (13.79), and provides us with the equation

$$\rho V(x) = -\frac{\gamma(V_x)^2}{V_{xx}} + \left(rx - \frac{1}{V_x} \right) V_x - \ln V_x, \quad (13.84)$$

where

$$\gamma = \frac{\alpha - r}{2\sigma^2}. \quad (13.85)$$

This is a nonlinear ordinary differential equation that appears to be quite difficult to solve. However, Karatzas, Lehoczky, Sethi, and Shreve (1986) used a change of variable that transforms (13.84) into a second-order, linear, ordinary differential equation. They assumed that the value function is strictly concave and, therefore, V_x is monotonically decreasing in x . This means that the function $c(\cdot)$ defined in (13.83) has an inverse $X(\cdot)$ such that (13.84) can be rewritten as

$$\rho V(X(c)) = -\frac{\gamma(U'(c))^2 X'(c)}{U''(c)} + (rX(c) - c)U'(c) + U(c), \quad (13.86)$$

where ' and '' denote, respectively, the first and second derivatives of functions with respect to their arguments. Differentiation with respect to c yields the intended second-order, linear ordinary differential equation

$$\gamma X''(c) = \left[(r - \rho - 2\gamma) \frac{U''(c)}{U'(c)} + \frac{\gamma U'''(c)}{U''(c)} \right] X'(c) + \left[\frac{U''(c)}{U'(c)} \right]^2 (rX(c) - c). \quad (13.87)$$

This equation has an explicit solution with three parameters to be determined; see Appendix A. After some calculations, one can determine these parameters, and obtain the solution of (13.84) as

$$V(x) = \frac{1}{\rho} \ln(\rho x) + \frac{r - \rho + \gamma}{\rho^2}, \quad x \geq 0. \quad (13.88)$$

13.3 Verify by direct substitution that the value function defined by (13.67) and (13.68) solves the HJB equation (13.66).

13.4 Verify by direct substitution that the value function in (13.88) solves the HJB equation (13.84).

13.5 Solve the consumption-investment problem (13.78) with the utility function

$$U(c) = c^\beta, \quad 0 < \beta < 1,$$

and $B = 0$.

13.6 Solve Exercise 13.5 when $U(c) = -c^\beta$ with $\beta < 0$ and $B = -\infty$.