Chapter 3

Stochastic Inventory Control

[1] In this chapter, we consider in much greater details certain dynamic inventory control problems of the type already encountered in section 1.3. In addition to the fact that this is a classical topic in stochastic control, we emphasize the following important idea. The dynamic programming algorithm is not only useful for computations, it is also a basic tool for the theoretical investigation of control problems. Using it, we prove here the optimality of the class of so-called base stock and (s, S)-policies for a classical formulation of the inventory management problem. References for this chapter are [Ber07, section 4.2], [Por02]. Numerous variations are possible starting from the basic inventory control problem, which aim at capturing different real-world situations.

3.1 Problem Formulation

In the dynamic inventory control problem (for a single product), the state x_k represents the level of inventory at the beginning of period k. We consider a slightly different formulation compared to section (1.3), and assume backlogging rather than lost sales if the demand at a given period cannot be met. Hence we assume now that x_k can take positive or negative values, which in the later case represents backlogged orders. Backlogged orders must be fulfilled first at the next period. Moreover, a unit penalty cost c_B is charged for each unit that is backlogged at each period. Using this device in fact allows us to ignore sales revenues and simply seek to minimize the expected cost of running the system. Each unit of positive leftover stock at the end of a period incurs a holding cost $c_H \geq 0$ in that period.

The time axis is divided as usual into discrete periods $k = 0, 1, \ldots, N$. At the beginning of period k, we observe the current inventory level x_k. We can decide how much additional product u_k to order, and we assume that the order is fulfilled immediately, i.e., is received in time to serve the demand w_k for that period. This demand is uncertain but we assume that it is stochastic, and more precisely that the w_k's are independent identically distributed (i.i.d.) and

[1] This version: September 19 2009

bounded random variables with a known distribution [2]. The state dynamics (1.1) are then

$$x_{k+1} = x_k + u_k - w_k, \; k = 0, 1, \ldots, N-1.$$

The (convex) holding and shortage cost for a period is of the form (see Fig. 3.1)

$$\mathcal{L}(x) = c_B\,(-x)^+ + c_H\,x^+,$$

where $x^+ = \max(0, x)$. In addition, there is a cost $c \geq 0$ per unit of ordered product. We assume $c_B > c$, so that it is not optimal to never order anything and simply accumulate backlog penalty costs. We also assume $c + c_H > 0$ (so at least one of these coefficients is positive). Finally, we assume that if the inventory level at period N is x, a *convex and nonnegative terminal cost* $v_N(x)$ is incurred [3]. We wish to minimize the total cost

$$E\left[\sum_{k=0}^{N-1} (cu_k + \mathcal{L}(x_k + u_k - w_k)) + v_N(x_N)\right].$$

Note that we decided, somewhat arbitrarily, that the shortage and holding costs are both charged at the end of each period (i.e., for the inventory left after the demand is realized). Applying the DP algorithm (1.12, 1.13), we obtain

$$J_N^*(x_N) = v_N(x_N)$$
$$J_k^*(x_k) = \min_{u_k \geq 0} \left\{ cu_k + L(x_k + u_k) + E_{w_k}\left[J_{k+1}^*(x_k + u_k - w_k\right]\right\}, 0 \leq k \leq N-1,$$

where we defined $L(y) = E_w[\mathcal{L}(y - w)]$, and E_w is the expectation with respect to the distribution of w_0 (recall that we assume the w_i's to be i.i.d.).

3.2 Proving Properties of the Value Function via DP and Backward Induction

For this problem, it is convenient to make a change of variable and define the new control to be the level of inventory after ordering $y_k = x_k + u_k$, with the corresponding constraint $y_k \geq x_k$. Define

$$G_k(y) = cy + L(y) + E_w[J_{k+1}^*(y - w)]. \tag{3.1}$$

Then we can rewrite the DP recursion as

$$J_k(x_k) = \left\{ \min_{y_k \geq x_k} G_k(y_k)\} \right\} - cx_k. \tag{3.2}$$

[2] Everything probably works with w_k dependent on x_k, u_k as in our basic model, but I haven't checked the details. Also, probably something weaker than the boundedness assumption works

[3] it might be possible to remove the nonnegativity assumption. I leave this for a future revision.

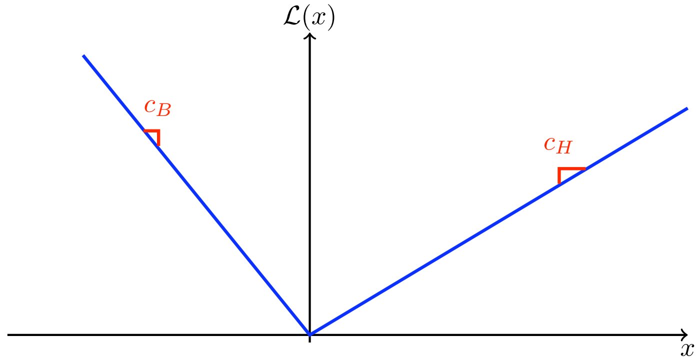

Figure 3.1: Immediate cost function in the dynamic inventory management problem.

Hence we see that the main issue becomes to understand the minimization problem for $G_k(y)$ over the interval $[x_k, \infty)$. Now assume for a moment that G_k has an unconstrained minimum at S_k

$$S_k \in \arg\min_{y \in \mathbb{R}} G_k(y).$$

Suppose also that for any such minimizer and $y \geq S_k$, G_k is nondecreasing (see Fig. 3.2). Then (in addition to guaranteeing that the set of minimizers is an interval), we see that the constrained minimization problem in this case has a simple solution: set $y = x_k$ if $x_k \geq S_k$, and $y = S_k$ if $x_k < S_k$. So we bring the inventory level back to S_k as soon as it drops below this level. In terms of the original control parameter, we would have the optimal policy of the form

$$\mu_k^*(x_k) = \begin{cases} S_k - x_k & \text{if } x_k < S_k \\ 0 & \text{if } x_k \geq S_k. \end{cases} \tag{3.3}$$

We will prove shortly that our conjecture on the shape of G_k is true, and that in fact something much stronger holds, namely that G_k is convex[4] and coercive, i.e.,

$$\lim_{|y| \to +\infty} G_k(y) = +\infty.$$

A decision rule of the form (3.3) is called a *base stock policy*. For practical use such a policy is very convenient[5]. We only need to compute in advance and record the set of scalars $S_0, S_1, \ldots, S_{N-1}$, called the *base stock levels*. Then at the beginning of period k, the inventory manager simply observes the inventory

[4]convexity could be guessed directly for J_k^*, as a result of partial minimization in a stochastic convex program.

[5]There are also computational advantages to know that there is such a simple parametrization of the optimal policy, which will become clear once we look at approximation methods, rollout and policy iteration.

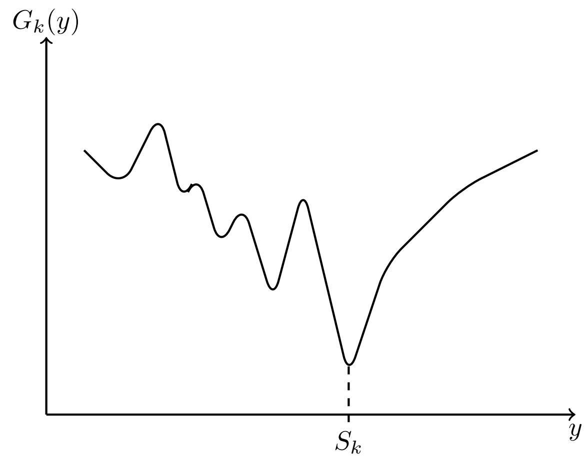

Figure 3.2: Structure of the function G_k guaranteeing the optimality of base stock policies with no fixed cost.

level x_k, does nothing if $x_k \geq S_k$, and brings it to level S_k otherwise by ordering the quantity $S_k - x_k$.

Our goal for the rest of the section is thus to prove that, for all $0 \leq k \leq N-1$, the function G_k is convex and coercive. This guarantees the existence of the base stock levels S_k and the optimality of the base stock policy (3.3). The method employed to do this is very important and must be remembered. We use the DP algorithm and a backward induction argument to prove a certain property of the optimal value function, in this case convexity. First note the following

Lemma 3.2.1. *If J^*_{k+1} is convex and nonnegative, then G_k is convex and coercive.*

Proof. This follows from the definition (3.1). First, by exercise 6 below, $y \to L(y)$ and $y \mapsto E_w[J^*_{k+1}(y-w)]$ are convex, so G_k is convex. Because J^*_{k+1} is nonnegative, we have $G_k(y) \geq cy + L(y)$. Then $\lim_{y\to+\infty} G_k(y) = +\infty$ because $c + c_H > 0$, and $\lim_{y\to-\infty} G_k(y) = +\infty$ because $c_B > c$ (and $|E[w]| < \infty$). $\square$

Exercise 6. Prove the result skipped above, namely, if $h : \mathbb{R} \to \mathbb{R}$ is a convex function, then $y \mapsto E_w[h(y-w)]$ is also convex.

Hence everything is proved if we show the following

Lemma 3.2.2. *The value function J^*_k is a convex and nonnegative function, for all $0 \leq k \leq N$.*

Proof. We proceed using backward induction. For $k = N$, we have $J^*_N(x) = v_N(x)$ and so the lemma is true by hypothesis. Suppose now that J^*_{k+1} is nonnegative and convex, we show that J_k^* is also nonnegative and convex. Now using lemma 3.2.1, we deduce the existence of S_k, a minimizer of G_k. As shown previously, the minimizer in the DP equation (3.2) is given by (3.3). We have then

$$J_k^*(x_k) = \begin{cases} c(S_k - x_k) + L(S_k) + E_w[J_{k+1}^*(S_k - w_k)] & \text{if } x_k < S_k \\ L(x_k) + E_w[J_{k+1}^*(x_k - w_k)] & \text{if } x_k \geq S_k. \end{cases}$$

or equivalently:

$$J_k^*(x_k) = \begin{cases} G_k(S_k) - cx_k & \text{if } x_k < S_k \\ G_k(x_k) - cx_k & \text{if } x_k \geq S_k. \end{cases}$$

Now the nonnegativity of J_k^* is immediate from the first expression since J_{k+1}^* and L are nonnegative. For the convexity, consider the second expression. The function which is constant equal to $G_k(S_k)$ for $x_k < S_k$ and then equal to $G_k(x_k)$ for $x_k \geq S_k$ is convex, since S_k is a minimizer of G_k (which is convex by lemma 3.2.1 and our induction hypothesis). Adding the linear function $x \mapsto -cx$ preserves convexity, so $J_k^*(x_k)$ is convex. This concludes the induction step. □

Theorem 3.2.3. *The base stock policy (3.3) is optimal at each period of the finite-horizon inventory control problem.*

Exercise 7. Convince yourself that we showed that the base stock policy (3.3) is indeed optimal by collecting the elements for the proof of the theorem 3.2.3.

Remark. In a problem that you encounter for the first time, it can be hard to find which property of the value function will be useful... Typically you can try to compute the solution for problems with a few periods (starting with 1; actually here the inventory control problem with one period is also important in management science, it is called the *newsvendor problem*). Then try to find a pattern that can be propagated through induction. The next section introduces a complication to the previous formulation that gives an additional example of the technique, and the problems and following chapters will also give you the opportunity to practice. It is useful to know some basic results and examples of problems with a nice structure. However, properties of the value function tend to be quite fragile, i.e., small changes in the assumptions of the problem can have a dramatic impact on the properties of the value function.

3.3 Adding a Positive Fixed Ordering Cost

Suppose now that ordering any quantity of product involves a fixed positive cost K in addition to the unit ordering cost. Thus, the cost for ordering a quantity u of product is 0 if no product is ordered and $K + cu$ if $u > 0$. The addition of this fixed cost significantly complicates the analysis of the problem, and in particular the value function is not convex any more in general. A more general notion, K-convexity, was introduced by Scarf [Sca60] to solve this problem. Intuitively, the base stock policy (3.3) should not be optimal in the presence of a fixed cost, which penalizes the frequent ordering of small quantities. Instead, we should delay new orders so that the fixed cost does not represent an exaggerated proportion of the total ordering cost. The optimal policy turns out to be characterized by two levels denoted S_k and s_k at each period, with $s_k < S_k$. It consist in waiting that the inventory level drops below s_k and then prescribes to bring the inventory back to level S_k. Such a policy is imaginatively called an (s, S) *policy*.

The DP recursion step (3.2) is now replaced by

$$J_k^*(x_k) = \min \left\{ G_k(x_k), \min_{y_k > x_k} (K + G_k(y_k)) \right\} - cx_k, \tag{3.4}$$

again using as new control variable $y_k = x_k + u_k$, the level of inventory after ordering. The first term in the minimization corresponds to not ordering any product. To understand this minimization problem, we plot $G_k(y)$, and look for the values y greater or equal to x_k. If there is a value of $G_k(y)$ that is smaller than $G_k(x_k)$ by more than K, then it is advantageous to move to that level of inventory. Referring to Fig. 3.3, we see that for $x_k \leq s_k$, we should set $y_k = S_k$, whereas for $x_k > s_k$, we should not change the inventory level because the function $G_k(y)$ never drops below $G_k(x_k)$ by more than K for $y \geq x_k$. We will see however that this function is not K-convex, hence K-convexity is only a *sufficient condition* guaranteeing the optimality of (s, S) policies. On the other hand, on Fig. 3.4 we have a function G_k for which an (s, S) policy is *not* optimal. Indeed if we have $\tilde{s}_k \geq x_k \geq t$ with $t > S_k$ and $\tilde{s}_k, \tilde{S}_k$ defined on the figure, then it is optimal to bring the level of inventory to $\tilde{S}_k$ instead of leaving it unchanged. This situation will be ruled out by K-convexity.

K-convexity is a generalization of convexity for functions of a single real variable. There are a number of equivalent characterizations of K-convexity, but perhaps the easiest way of seeing that the functions on Fig. 3.3 and 3.4 are not K-convex is by using the following definition

Definition 3.3.1. A function $f : \mathbb{R} \to \mathbb{R}$ is K-convex, with $K \geq 0$, if for each $y \leq y'$, $0 \leq \theta \leq 1$, we have

$$f(\theta y + (1-\theta)y') \leq \theta f(y) + (1-\theta)(K + f(y')).$$

Hence clearly a 0-convex function is synonymous with a convex function. For a function to be K-convex, it must lie below the line segment connecting $(y, f(y))$ and $(y', f(y') + K)$ on the interval $[y, y']$. Such a function can be discontinuous, but the jumps at the discontinuity point must necessarily be downwards and cannot be too large. A K-convex function for $K > 0$ need not be convex or even quasi-convex, and can have several local minima. The following property is immediate from the definition.

Lemma 3.3.1. *If f is K-convex, $y < y'$, and $f(y) = K + f(y')$, then $f(z) \leq K + f(y')$ for all $z \in [y, y']$.*

Exercise 8. Prove Lemma 3.3.1.

A K-convex function can cross the value $K + f(y)$ at most once on $(-\infty, y)$. We see immediately then that the function on Fig. 3.3 is not K-convex because it does not remain under the horizontal dotted line $f(y') + K$ on the interval $[y, y']$ shown on the figure. The same idea applies to the function of Fig. 3.4. Here are some additional properties of K-convex functions.

Lemma 3.3.2. *1. A function $f : \mathbb{R} \to \mathbb{R}$ is K-convex iff*

$$K + f(y + a) \geq f(y) + \frac{a}{b}[f(y) - f(y - b)], \text{ for all } x \in \mathbb{R}, a \geq 0, b > 0.$$

2. If f is differentiable, then f is K-convex iff

$$K + f(y) \geq f(x) + f'(x)(y - x), \text{ for all } x \leq y.$$

3. If f_1 and f_2 are K- and L-convex ($K, L \geq 0$), and $\alpha, \beta > 0$, then $\alpha f_1 + \beta f_2$ is $(\alpha K + \beta L)$-convex.

4. If f is K-convex and w is a random variable, then $E_w[f(y - w)]$ is also K-convex, provided that $E_w[|f(y - w)|] < \infty$ for all y.

Exercise 9. Prove Lemma 3.3.2.

Next we show that if G_k is K-convex, continuous, and coercive, then an (s, S) policy is optimal at stage k.

Lemma 3.3.3. *If f is a continuous and coercive K-convex function, then there exist scalars $s \leq S$ such that*

1. *S minimizes f: $f(S) \leq f(y)$, $\forall y \in \mathbb{R}$.*
2. *$f(S) + K = f(s)$ and $f(y) > f(s)$ for all $y < s$.*
3. *$f(y)$ is a decreasing function on $(-\infty, s)$.*
4. *$f(y) \leq f(y') + K$ for all $s \leq y \leq y'$.*

Proof. S exists because f is continuous and coercive. Define s to be the smallest number z in $\mathbb{R}$ such that $z \leq S$ and $f(S) + K = f(z)$. The rest follows more or less directly from lemma 3.3.1, it is best to draw a picture. □

Exercise 10. Use lemma 3.3.3 to show that if G_k is continuous, K-convex and coercive, with K equal to the fixed ordering cost, then an (s, S) policy is optimal at stage k. Give a characterization of the inventory levels s and S in terms of the value of G_k at these points.

The final step consists in showing that G_k is K-convex, continuous, and coercive, for $k = 0, \ldots, N - 1$. This implies the optimality of an (s, S) policy at all stages by lemma 3.3.3 and exercise 10. As you have probably guessed by now, we show the crucial properties of G_k by backward induction. The proof follows the argument of section 3.2 with convexity replaced by K-convexity.

Lemma 3.3.4. *If J^*_{k+1} is K-convex, nonnegative and continuous, then G_k is K-convex, coercive, and continuous.*

Proof. K-convexity and coercivity of G_k follow as in lemma 3.2.1, using the properties in lemma 3.3.2. Continuity of G_k uses the assumption that w is bounded to get that $L(y)$ and $E_w[J^*_{k+1}(y-w)]$ are continuous.[6] □

Hence it is now sufficient to show the following result.

Lemma 3.3.5. *The value function J^*_k is a K-convex, nonnegative and continuous function, for all $0 \le k \le N$.*

Proof. At stage N, we have $J^*_N(x_N) = v_N(x_N)$ convex and nonnegative, so the result is true by hypothesis. Now assume that J^*_{k+1} is K-convex, nonnegative and continuous. This implies that G_k is K-convex, coercive and continuous by lemma 3.3.4. Next, J^*_k is defined by (3.4). We deduce that J^*_k has the following form

$$J^*_k(x) = \begin{cases} K + G_k(S_k) - cx, & \text{for } x \le s_k, \\ G_k(x) - cx, & \text{for } x \ge s_k, \end{cases}$$

where s_k, S_k are defined as in lemma 3.3.3. J^*_k is continuous, by definition of s_k and since G_k is continuous. The nonnegativity of J^*_k follows say from the definition of J^*_k as cost-to-go, where all costs are nonnegative. We want to show that J^*_k is K-convex. Equivalently since $x \mapsto cx$ is 0-convex and by lemma 3.3.2, we want to show that $\tilde{G}_k$ defined by $\tilde{G}_k(x) = J^*_k(x) + cx$ is K-convex. We consider definition 3.3.1. If $s_k \le y \le y'$ then $\tilde{G}_k = G_k$ with G_k K-convex so the K-convexity inequality is satisfied. If $y \le y' \le s_k$ then $\tilde{G}_k$ is constant (equal to $G_k(s_k) = K + G_k(S_k)$) so again the K-convexity inequality is satisfied. The remaining case is $y < s_k < y'$. Consider $z \in [y, y']$. We need to show that $\tilde{G}_k(z)$ is below the line segment connecting $(y, G_k(s_k))$ (since $\tilde{G}_k(y) = G_k(s_k)$) and $(y', G_k(y') + K)$ (since $\tilde{G}_k(y') = G_k(y')$). First note that $K + G_k(y') \ge K + G_k(S_k) = \tilde{G}_k(y)$ so this line segment is "increasing". Then if $y \le z \le s_k$ this is clear because $\tilde{G}_k$ is constant on that interval. If $s_k \le z \le y'$ we first have by K-convexity on $[s_k, y]$ that $\tilde{G}_k(z)$ is below the line segment connecting $(s_k, G_k(s_k))$ and $(y', G_k(y') + K)$. Then it is not hard to see (draw it!) that this later line is itself below the line of interest. This concludes the proof of the K-convexity of J^*_k and the induction step. □

3.4 Practice Problems

Problem 3.4.1. Do all the exercises found in the chapter.

[6] a better argument and assumption on w can probably be made here.

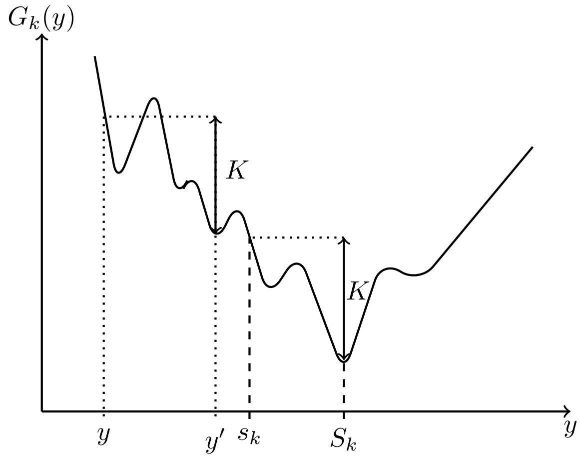

Figure 3.3: Structure of the function G_k guaranteeing the optimality of an (s, S) policy (although this function is not K-convex).

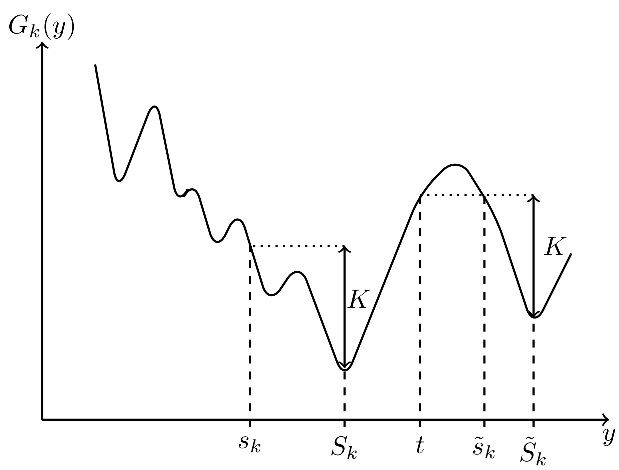

Figure 3.4: A function G_k for which an (s, S) policy is *not* optimal.