Manejo de Inventario

David Peña Peralta

2025-07-11

Table of contents

1 Introducción

Dentro del area del Control Estocástico, una de los problemas más conocidos son los problemas de inventario. Donde se presenta una bodega con capacidad máxima K. Cada estapa se extrae una cantidad de mercancía, la que denotaremos como la demanda D_t, y se solicita una cantidad del producto a_t, obteniendo finalmente el nivel de inventario X_t (ver Sutton and Barto (2018)). En general se busca minimizar los costos de la bodega (costos por almacenamiento, costos por pérdida, entre otros).

2 Formulación del Proceso de Decisión de Markov.

Para nuestro problema consideraremos un supermercado, centrado en uno de sus pasillos. Suponiendo que en un pasillo se almacena un solo tipo de producto. Definiremos a K la cantidad máxima de producto en el pasillo, X_t a la cantidad del producto disponible para la venta (o la cantidad de producto en el pasillo). Nuestra demanda, o producto solicitado, será denotado por D_t y se considerará una colección de v.a i.i.d. Finalmente, la cantidad recolocada en el pasillo, o producto pedido, será denotada por a_t. Entonces, nuestro conjunto de estados $\mathcal{S}$ está dado por el siguiente conjunto

$$\mathcal{S} = \{s \in \mathbb{Z}^+ : 0 \leq s \leq K\}. \tag{2.1}$$

Nuestro conjunto de acciones $\mathcal{A} = \mathcal{S} = \mathbb{Z}^+$, y para $x \in \mathcal{S}$ nuestro conjunto de acciones admisibles esta dado por

$$\mathcal{A}(x) = \{a \in \mathcal{A} : 0 \leq a \leq K - x\}.$$

3 Dinámica del Modelo.

Recordando la fórmula para nuestro modelo.

$$X_{t+1} = f(X_t, a_t), f : \mathcal{S} \times \mathcal{A} \to \mathcal{S}.$$

Entonces el modelo que usaremos esta dado por

$$X_{t+1} = (X_t + a_t - \eta X_t - D_{t+1})^+, \tag{3.1}$$

donde a_t es la cantidad de producto recolodado al final del dia t, η es el factor descomposición, D_t es la demanda del prodcucto en la dia t y $(\cdot)^+ = \max\{\cdot, 0\}$.

4 Descripción y Justificación del Modelo.

El modelo Equation 3.1 pretende responder a la pregunta que denota el modelo ¿Cuánto producto tendré disponible al dia siguiente?. Lo anterior menciona que nuestras etapas $t \in \mathcal{T} = \{t \in \mathbb{Z}^+ : t \leq T, T \in \mathbb{N}\}$ representaran los dias dentro de un periodo T, t hace referencia al dia actual, y $t+1$ al dia siguiente. Entonces el modelo general esta dado por

$$X_{t+1} = (\text{Today} + \text{In}_t - \text{Out}_{t+1})^+.$$

Esto es, la parte positiva del producto que hay "hoy", es decir, X_t. A eso le agregaremos el producto que entrará hoy al final del dia, en nuestro modelo solo habrá ingreso de producto mediante solicitud (En este caso no consideramos un almacenimiento dentro del supermercado), entonces In_t esta dado por nuestras acciones $\text{In}_t = a_t$.

La parte que saldrá consta de dos elementos. En general consideramos la cantidad de producto que se compró en el dia t. Sin embargo, desconocemos la cantidad requerida, haciendo referencia al dia siguiente. Por lo tanto la demanda está representada por D_{t+1}, la cantidad de producto requerida al dia siguiente. En nuestro modelo también consideramos la salida de producto por considerarse producto no apto para la venta. Entonces

$$\text{Out}_t = D_{t+1} + N_t(X_t).$$

Bajo de la suposición que todos los productos poseen el mismo tiempo de vida con periodos de vida distintos supondremos que cada dia, al final, se retira un factor con respecto a la cantidad actual de producto.

$$N_t = \eta X_t$$

$$\text{Out}_t = D_{t+1} + \eta X_t$$

Finalmente, nos queda definir la función de costo, en nuestro modelo será la ganancia. Al considerar un periodo finito tenemos que la ganancia total G esta dada por

$$G(x_0, \pi) = \sum_{t=0}^{T} G_t(X_t, a_t), X_0 = x_0, X_{t+1} = f(X_t, a_t).$$

donde π es una politica, $\pi = (a_0, a_1, \dots, a_{N-1})$. y G_t es la ganancia por etapa, en nuestro caso

$$G_t(x,a) = P_V \min\{x+a, D_t\} - P_S(a - \mathcal{I}_{t=0}x),$$

notemos que en el dia $a = 0$ y $D_0 = 0$, entonces $G_0(x,a) = -P_S x$ donde C es el costo unitario por tener el producto al inicio. Notemos que D_t es una variable aleatoria, entonces la función de valor por estado es la siguiente

$$V^\pi(s) = E[G(s,\pi)]$$

Teniendo que la ecuación de Bellman para nuestra función de valor es

$$V^\pi(s) = \sum_a \pi(a \mid s) \sum_{s'} \mathcal{P}[s' \mid s, a][R(s', a, s) + \gamma V^\pi(s')]$$

5 Justificación de las acciones.

Ya comentamos que nuestras acciones, serán la cantidad de producto que vamos a solicitar. Entonces nuestras acciones serán números enteros las acciones serán ejecutadas de forma instantea. El conjunto de acciones está dado por

$$\mathcal{A} = \{z \in \mathbb{Z}^+ : z \leq K\}.$$

y para cada $x \in \mathcal{S}$, obtenemos el conjunto de acciones admisibles.

$$\mathcal{A}(x) = \{z \in \mathbb{Z}^+ : z \leq K - x\}.$$

References

Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction.* Cambridge, MA, USA: A Bradford Book.