Fitted Q-Iteration: Estimación de niveles de presas

Erika Rubí Jiménez Guzmán

2028-04-04

Table of contents

Resumen

Aquí pondré un resumen de mi tesis, que se llama “Fitted Q-Iteration: Estimación de niveles de presas”.

Introduction

Esta será la introducción.

ejemplo de una cita Knuth (1984)

Part I

Preliminares

1 Inteligencia Artificial

La IA es...

1.1 Machine Learning (Aprendizaje Automático)

Disciplina dentro de la IA... Tipos principales

Definition 1.1 (Aprendizaje Supervisado)**.** Se entrena con datos etiquetados...

Definition 1.2 (Aprendizaje No Supervisado)**.** Se entrena con datos no etiquetados...

Definition 1.3 (Aprendizaje por Refuerzo)**.** El agente aprende a través de la interacción con el entorno...

2 Q-Learning

Algoritmo de aprendizaje por refuerzo...

Definition 2.1 (Fitted Q-Iteration)**.** Algoritmo Offline (aprende analizando el historial de experiencias)...

Part II

Modelo Matemático

3 Modelo Matemático para Optimizar la Operación de un Sistema Hidroeléctrico.

El objetivo de este capítulo es presentar una formulación rigurosa del problema de optimización de la operación de un sistema hidroeléctrico como un Proceso de Decisión de Markov (MDP), junto con una descripción detallada del algoritmo de Fitted Q-Iteration (FQI) utilizado para aproximar la política óptima. Se enfatizará la conexión entre los elementos matemáticos formales y su implementación numérica, así como la validación operativa mediante simulación en lazo cerrado.

3.1 Formulación como Proceso de Decisión de Markov (MDP)

3.1.1 Estructura Temporal

Se considera un horizonte temporal de operación compuesto por $T = 50$ años históricos, en donde cada uno de ellos se dividieron en $q = 24$ quincenas. Con el fin de reducir la dimensionalidad temporal y poder capturar la estacionalidad, se agruparon las quincenas en $M = 6$ etapas hidrológicas mediante la siguiente función de agregación $\mu : \{0, \dots, q-1\} \to \{0, \dots, M-1\}$ definida por

$$\mu(q) = \begin{cases} 0, & 0 \leq q < 10 \\ 1, & 10 \leq q < 14 \\ 2, & 14 \leq q < 16 \\ 3, & 16 \leq q < 18 \\ 4, & 18 \leq q < 20 \\ 5, & 20 \leq q < 23 \end{cases}$$

así, la variable $m = \mu(q)$ representará la etapa hidrológica correspondiente a la quincena q. El índice temporal completo se denotará por como $\tau = (t, q)$ con $t \in \{0, \dots, T-1\}$ y $q \in \{0, \dots, q-1\}$.

3.1.2 Espacio de estados y operador de discretización

El sistema está compuesto por dos presas: La Angostura ($i = 1$) y Malpaso ($i = 2$). El volumen de almacenamiento de cada una de ellas en la quincena q del año t se representará por $V_{i,t,q} \in \mathbb{R}_{\geq 0}$ (Mm^3).\\ Dado que los datos históricos se utilizan para construir un conjunto de transión empírica, se introduce un operador de discretización uniforme del espacio de estados mediante el parámetro $\Delta = 600\ Mm^3$. Así, el espacio de estados se define como $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \{0, ..., 5\} = \{s = (s_1, s_2, m)\}$ en donde s_i representa el volumen discretizado de la presa i, tal estado discreto se obtiene mediante el operador de discretización $\mathcal{D} : \mathbb{R} \times \mathbb{N} \longrightarrow \{0, \dots, N-1\}$ dado por

$$s_i = \mathcal{D}(V_{i,t,q}, N_i) = \min\left(\max\left(\left\lfloor \frac{V_{i,t,q}}{\Delta} + \frac{1}{2} \right\rfloor, 0\right), N_i - 1\right),$$

donde $N_1 = 27$ y $N_2 = 17$ son las capacidades discretas máximas de cada presa.\\

3.1.3 Espacio de acciones y restricciones operativas dependientes de la etapa

- Conjuntos $\mathcal{K}_{i,m}$, vector de unidades u_m, y espacio producto $\mathcal{A}_s = \mathcal{K}_{1,m} \times \mathcal{K}_{2,m}$.
- Comentar la dependencia estacional y la estructura de retículos enteros.

3.1.4 Dinámica de transición y construcción del kernel empírico

- Ecuación de balance hídrico continua y proyección discreta.
- **Elemento matemático sugerido:** *Definición X.2 (Kernel empírico-determinista)* $\mathcal{P}(s'|s, a) = \mathbb{I}\{s' = \Phi(s, a, I_\tau)\}$ y discusión sobre muestreo histórico como aproximación de la distribución hidrológica real.

3.2 Función de recompensa y criterio de optimalidad

3.2.1 Modelo hidráulico-energético linealizado

- Aproximación de la altura $H_i(s_i, s_i')$ y fórmula de generación $E_i = \alpha H_i k_i u_m$.
- Justificación física y rango de validez del *surrogate* lineal.

3.2.2 Estructura de penalizaciones y manejo de restricciones blandas

- Definición de $\Pi_{\text{spill}}, \Pi_{\text{def}}, \Pi_{\text{CG}}, \Pi_{\text{crit}}$ y operador $(\cdot)^+$.
- **Elemento matemático sugerido:** *Observación X.1* sobre la equivalencia entre penalizaciones cuadráticas/lineales y métodos de barrera interior en optimización restringida.

3.2.3 Ecuación de optimalidad de Bellman descontada

- Formulación de $Q^*(s,a) = \mathcal{R}(s,a,s') + \gamma \max_{a'} Q^*(s',a')$.
- Declaración explícita del criterio de optimalidad esperado descontado $J(\pi)$.

3.3 Aproximación numérica mediante Fitted Q-Iteration (FQI)

3.3.1 Motivación: maldición de la dimensionalidad y aprendizaje por refuerzo fuera de política

- Justificación de no usar iteración de política clásica o programación dinámica exacta.
- Referencias a Ernst et al. (2005), Antos et al. (2008).

3.3.2 Iteración de Bellman empírica y estabilización numérica

- Cálculo de $y_n^{(j)}$, operador de truncamiento $\mathrm{clip}(\cdot)$, y manejo de divergencias en métodos *off-policy*.
- **Elemento matemático sugerido:** *Algoritmo X.1* en pseudocódigo estructurado (estilo ACM/IEEE).

3.3.3 Aproximación funcional por regresión no paramétrica

- Problema de mínimos cuadrados empírico sobre $\mathcal{F}$, hiperparámetros (árboles, profundidad, *min_samples_leaf*), y justificación de *Extra Trees* (reducción de varianza, sesgo controlado).
- Comentario sobre universalidad y capacidad de aproximación de bosques aleatorios.

3.3.4 Criterios de convergencia y tolerancia computacional

- Norma media en conjunto fijo $\mathcal{S}_{\text{test}}$, umbral $\varepsilon = 10^{-3}$, y detección temprana de estancamiento.
- **Elemento matemático sugerido:** *Criterio de parada X.1* con justificación numérica (estabilidad de punto fijo aproximado).

3.4 Protocolo de simulación en lazo cerrado y evaluación operativa

3.4.1 Extracción de la política greedy óptima

- Definición de $\pi^*(s) \in \arg\max_{a \in \mathcal{A}_s} \hat{Q}^*(s, a)$ y discusión sobre no unicidad y desempate determinista.

3.4.2 Dinámica forward y proyección sobre restricciones físicas

- Simulación sobre series históricas, proyección $\min\{\max\{\tilde{V}_i, 0\}, N_i \Delta V\}$, y manejo de fronteras anuales.
- Comentario sobre desacople entre entrenamiento (penalizaciones blandas) y evaluación (restricciones duras).

3.4.3 Métricas de desempeño y validación estadística

- Energía esperada por reserva, derrames acumulados, violaciones de curvas guía, y análisis de varianza interanual.
- **Elemento sugerido:** *Tabla X.1* resumen de parámetros operativos y *Figura X.1* trayectoria típica de volúmenes y energía.

3.5 Análisis teórico y propiedades del modelo

3.5.1 Contractividad del operador de Bellman en norma ℓ_∞

- **Lema X.1:** $\|\mathcal{T} Q_1 - \mathcal{T} Q_2\|_\infty \leq \gamma \|Q_1 - Q_2\|_\infty$.
- Breve demostración usando la propiedad $\max - \max \leq \max(|\cdot|)$ y citando Puterman (1994, Thm. 6.2.3).

3.5.2 Cotas de error de aproximación y complejidad muestral

- Descomposición del error: $\|\hat{Q}^* - Q^*\|_\infty \leq \frac{2\gamma}{(1-\gamma)^2} \epsilon_{\text{approx}} + \mathcal{O}(N^{-1/2})$.
- Comentario sobre trade-off sesgo-varianza en $\mathcal{F}$ y efecto del tamaño de muestra histórico.

3.5.3 Alineación con el marco axiomático de Puterman (1994)

- **Tabla X.2** correspondencia código $\leftrightarrow$ teoría MDP.
- Discusión sobre validez de supuestos: horizonte infinito, descuento, espacios finitos, kernel empírico como aproximación de $\mathcal{P}$.

3.6 Limitaciones, extensiones y trabajo futuro

- Limitaciones: discretización fija, modelo lineal de altura, ausencia de incertidumbre explícita en afluentes, penalizaciones estáticas.
- Extensiones: curvas cota-volumen reales, kernels estocásticos (ARIMA, copulas), FQI con regularización, métodos de política primal-dual para restricciones duras.
- Aplicación a otros sistemas multirreservorio y escalabilidad computacional.

3.7 Conclusiones del capítulo

- Síntesis de contribuciones matemáticas y numéricas.
- Reafirmación de la coherencia entre formulación teórica, implementación algorítmica y validación operativa.
- Declaración de cierre alineada con los objetivos de la tesis.

4

5

Referencias

Knuth, Donald E. 1984. “Literate Programming.” *Comput. J.* (USA) 27 (2): 97–111. https://doi.org/10.1093/comjnl/27.2.97.