Research School for Operations Management and Logistics

Zero-shot Generalization in Inventory Management: Train, then Estimate and Decide

Tarkan Temizoz

Deploying deep reinforcement learning (DRL) algorithms in real-world inventory management presents challenges, including dynamic operational environments and uncertainty regarding problem parameters, e.g. demand distributions. These challenges underline a significant research gap, indicating the need for a unifying framework to model and solve sequential decision-making problems under parameter uncertainty. We address this gap by exploring an underdeveloped area of DRL for inventory management: training generally capable agents (GCAs) under zero-shot generalization. In our context, this approach refers to advanced DRL policies that perform well across a broad range of problem instances and can be directly applied to new instances with unknown parameters without explicit retraining.

We propose a unifying Super-Markov Decision Process formulation and the Train, then Estimate and Decide (TED) framework to train and deploy a GCA tailored to inventory management applications. The TED framework consists of three phases: training a GCA on varied problem instances, continuously estimating problem parameters during deployment, and making decisions based on these estimates. Applied to periodic review inventory problems with lost sales, cyclic demand patterns, and stochastic lead times, our trained agent, the Generally Capable Lost Sales Network (GC-LSN) consistently outperforms well-known traditional policies when problem parameters are known. Moreover, under conditions where demand and/or lead time distributions are initially unknown and must be estimated, GC-LSN demonstrates superior performance compared to state-of-the-art online learning methods.