Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models

Linwei Xin
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

David A. Goldberg
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Inventory models with lost sales and large lead times have traditionally been considered intractable, due to the curse of dimensionality which arises from having to track the set of orders placed but not yet received (i.e. pipeline vector). Recently, Goldberg et al. (2012) laid the foundations for a new approach to solving these models, by proving that as the lead time grows large (with the other problem parameters fixed), a simple constant-order policy (proposed earlier by Reiman (2004)) is asymptotically optimal. This was quite surprising, as it is exactly this setting (i.e. large lead times) that was previously believed intractable. However, the bounds proven there are impractical, requiring the lead time to be very large before the constant-order policy becomes nearly optimal, e.g. requiring a lead time which is \( \Omega(\epsilon^{-2}) \) to ensure a \((1 + \epsilon)\)-approximation guarantee, and involving a massive prefactor. The authors note that numerical experiments of Zipkin (2008b) suggest that the constant-order policy performs quite well even for small lead times, and pose closing this gap (thus making the results practical) as an open problem.

In this work, we make significant progress towards resolving this open problem and closing this gap. In particular, for the infinite-horizon variant of the finite-horizon problem considered by Goldberg et al. (2012), we prove that the optimality gap of the same constant-order policy actually converges \textit{exponentially fast} to zero, i.e. we prove that a lead time which is \( O(\log(\epsilon^{-1})) \) suffices to ensure a \((1 + \epsilon)\)-approximation guarantee. We demonstrate that the corresponding rate of exponential decay is at least as fast as the exponential rate of convergence of the expected waiting time in a related single-server queue to its steady-state value. We also derive simple and explicit bounds for the optimality gap. For the special case of exponentially distributed demand, we further compute all expressions appearing in our bound in closed form, and numerically evaluate them, demonstrating good performance for a wide range of parameter values. Our main proof technique combines convexity arguments with ideas from queueing theory.

\textit{Key words}: inventory, lost-sales, constant-order policy, lead time, asymptotic optimality, steady-state.
1. Introduction

It is well-known that there is a fundamental dichotomy in the theory of inventory models, depending on the fate of unmet demand. If unmet demand remains in the system and can be met at a later time, we say the system exhibits backlogged demand; if unmet demand is lost to the system, we say the system exhibits lost sales. Which of these assumptions is appropriate depends heavily on the application of interest. For example, in many retail applications one must manage an inventory in a competitive environment, i.e. demand can in principle be met by a competing supplier, making lost sales a more appropriate assumption. Indeed, as pointed out in Bijvank and Vis (2011), recent studies have shown that retailers across many sectors lose over 75% of the potential demand which they cannot satisfy immediately, and we refer the interested reader to Gruen, Corsten and Bharadway (2002), and Verhoef and Sloot (2006) for further details.

A second important feature of many inventory models, intimately related to the above dichotomy, is that of positive lead times, i.e. settings in which there is a multi-period delay between when an order for more inventory is placed and when that order is received. In principle, this feature leads to an enlarged state-space (growing linearly with the lead time), to track all orders already placed but not yet received, i.e. the pipeline vector. It is a classical result, indeed one of the foundational results of the field, that models with backlogged demand remain tractable even in the presence of positive lead times. Namely, it can be proven that a so-called base-stock (i.e. order-up-to) policy, based only on the total inventory position (i.e. sum of the current inventory and all orders in the pipeline vector), is optimal in this setting (cf. Scarf (1960), Iglehart (1963), Veinott (1966)). Intuitively, this follows from the fact that when demand is backlogged, inventory is a linear function of orders placed and past demands, along with certain convexity arguments. However, it is known that such simple policies are no longer optimal for models with lost sales and positive lead times (cf. Karlin and Scarf (1958)). For over fifty years, inventory models with lost sales and positive lead times were