Top
DOI 10.1007/s11750-010-0160-x

ORIGINAL PAPER

Newsvendor games: convex optimization of centralized inventory operations

Samuel Burer · Moshe Dror

Received: 14 January 2010 / Accepted: 9 September 2010

Abstract A finite set of outlets with randomly fluctuating demands bands together to reduce costs by buying, storing and distributing their inventory jointly. This is termed inventory centralization and is a type of risk pooling. The expected centralization cost can be lowered even further, without disrupting the demand behavior at individual outlets, by inducing the outlets to correlate their individual demands. Given that the outlets' demands are normally distributed, the lowering of the centralized cost corresponds to a semidefinite optimization problem. This paper establishes a closed-form optimal solution of the semidefinite program and a fair allocation of the centralized cost at optimality.

Keywords Inventory centralization · Risk pooling · Cooperative game theory · Convex optimization

Mathematics Subject Classification (2000) 90B05 · 91A12 · 90C22 · 90C25

1 Introduction

Centralized inventory allows a coalition of outlets to reduce their expected inventory costs via the dampened variation in demand experienced at the warehouse. By appropriately correlating their individual demands further, additional savings can be realized. In order to provide minimal incentives for the outlets to maintain the centralized inventory arrangement, the expected costs should be allocated among the outlets so that no subset can be rewarded by breaking off from the coalition.

Research of S. Burer was supported in part by NSF Grants CCR-0203426 and CCF-0545514.

S. Burer (✉)
Department of Management Sciences, University of Iowa, Iowa City, IA 52242-1994, USA
e-mail: samuel-burer@uiowa.edu

M. Dror
MIS Department, The University of Arizona, Tucson, AZ, USA

Published online: 08 October 2010

There are many real-life examples where companies attempt to reduce inventory and manufacturing costs by manipulating correlated demand sources, or inducing demand patterns that balance each other in an average sense (Gerchak and Gupta 1991).

The hockey-stick phenomenon occurs in "just about every company" (Chase and Aquilano 1992). This term refers to the practice of shipping most of the demand from a factory during the last week or two of a fiscal period such as a quarter. This intensive correlation of demand at the end of the period has very bad effects on a manufacturing or order-fulfillment operation. Most of the work falls in a short time span, which increases chances for errors, reduces quality control, and causes missed shipping deadlines. Some companies have gone to great lengths to reduce this effect. For instance, in the 1980s, Digital Equipment Corporation's US operations staggered fiscal years-ends for its regions to end in each of the 12 months, so that the quarter-end rush to ship computing equipment would be averaged out over the calendar year. They took this extreme measure to reduce the overlap in deadlines and to simplify and improve their shipping schedules. Such an effort shows the importance companies place on controlling the overall demand pattern.

Lee et al. (1997) observed manufacturers going to great lengths to smooth out variances. The problem these authors discuss is the amplification of demand variance (the bull whip effect) as forecasts flow from lower levels to higher ones in a distribution chain. Timing of forecasts and their frequency may cause an inventory system to oscillate between shortage and excess inventory in a multiperiod setting. Some of the techniques they observed companies using are designed specifically to produce a negative correlation between demand points. These include staggering the times for receipt of forecasts and shipping orders at different times.

The literature of utility management, such as power plant grids, supplies examples of attempts to decouple or decorrelate demand. For instance, Ruusunen et al. (1991) discuss a multiperiod analysis of a group of interconnected utilities. An attempt is made to negatively correlate demand by using off-peak price reductions at different utilities.

Another example: a cola or beer bottling plant may attempt to stagger the peak demand at different service regions (e.g., via regional price promotions) so as to reduce the peak production demand at the plant. The mean demand at each region will hopefully remain the same (maintaining the same volume or market share in a highly competitive environment is considered a marketing and operational success). However, manipulating the buying patterns among the different regions appropriately might reduce the peak production pressure on the plant. For benefits of pooling in production-inventory systems see Benjaafar et al. (2005).

It is clear that companies understand well the benefits of manipulating correlations between demand points. Each company will determine its own strategy regarding the correlations based on the characteristics of its market. The challenge is to explain it analytically and to provide computational strategies that determine an optimal correlation solution and allocate costs fairly.

In this paper, we restrict our attention to outlets with single-period, normally distributed, correlated individual demands and identical linear holding and penalty costs. We assume an infinite horizon setting with risk neutrality for costs and benefits. In the context of inventory centralization, this *newsvendor* setting has first been studied by Eppen (1979), who establishes the potential for savings by showing that the expected cost of centralization can be expressed as a constant multiple of the standard deviation at the warehouse. (Another related point of view found in the literature is that of *capacity centralization*.)

Many authors have extended Eppen's results to other models with varying assumptions. For instance, Gerchak and He (2003) examine the relation between the benefits of risk pooling and the variability of demand and provide an example where increased variability of individual demands reduces the benefits of risk pooling. Benjaafar et al. (2005) examine the benefits of risk pooling in production-inventory systems with endogenously generated lead times. The effect of centralization on expected profit in a multi-location newsvendor setting is summarized by Cherikh (2000). Eppen's results have also been used in the literature of supply chain planning; see for example the paper of Shen and Coullard (2003), who utilize Eppen's risk pooling in a location-inventory model under the assumption of independent normal demands. For early work on cooperative games in the context of inventory centralization, see Gerchak and Gupta (1991), Hartman (1994), and Hartman and Dror (1996).

The models of Eppen and others assume that the joint distribution of demand is given, i.e., is an input to the model. In contrast, Hartman and Dror (2003) extend Eppen's idea by considering the following question, which matches the spirit of the examples described above: how far could the centralization cost be lowered if it were possible to coordinate the pairwise correlations between outlets to further reduce the variance at the warehouse? Hartman and Dror show that this minimization of the centralization cost leads to a semidefinite optimization problem.

The first broad goal of this paper is to study the structure of the semidefinite optimization problem introduced by Hartman and Dror. By extending some results from a branch of statistics called *minimum trace factor analysis*, we establish an easily computable, closed-form optimal solution. Further, we discuss how the structure of the optimal solution provides insight into real-world strategies for demand smoothing.

A second broad goal of this paper is to provide a fair allocation of costs to the outlets, where the precise meaning of "fair" is as described by Hartman and Dror (1996) (see Sect. 2). We examine the case where the cost allocations are allowed to be negative (i.e., an outlet is given money to participate because its presence provides a significant benefit to the grand coalition), as well as the more typical case when all cost allocations are nonnegative. Our key result is: if an optimal inventory centralization has been achieved, i.e., via the semidefinite optimization, then a fair cost allocation (either nonnegative or unrestricted) can be stated in a simple, closed form.

2 Elements of the model

Let $U = \{1, \dots, n\}$ be a set of outlets. Given a fixed duration time period, we assume that demand at the outlets is jointly normal. In particular, each outlet $i \in U$ is normally distributed with mean $\mu_i > 0$ and standard deviation $\sigma_i > 0$, and correlations ρ_{ij} are encoded in the correlation matrix R. We consider a static infinite horizon setting (i.e., an infinite number of time periods) where expected inventory cost is the metric. Inventory cost is composed of the holding cost h (per unit held) and the penalty cost p (per unit out of stock); we assume these numbers are constant for any coalition $S \subseteq U$ of outlets that bands together to centralize inventory.

We assume throughout this paper that the vectors μ and σ are fixed, i.e., are inputs to the model. On the other hand, with regards to the correlation matrix R, we will investigate the effect that changing R has on the model.

The expected inventory cost for an assembly S is proportional to the standard deviation of their joint distribution (Eppen 1979; Hartman et al. 2000), where the proportionality constant depends only on p, h, and the standard normal distribution and hence can be scaled to 1. Thus, with respect to R, the expected cost for any S can be represented by

$$c_R(S) = \sqrt{\sum_{i \in S} \sum_{j \in S} \sigma_{ij}} = \sqrt{\sum_{i \in S} \sum_{j \in S} \sigma_i \sigma_j \rho_{ij}},$$

where $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$ represent the covariances of the joint distribution of demand for S. In particular, the cost for a singleton outlet i is $c_R(\{i\}) = \sigma_i$.

2.1 Optimizing centralized inventory costs

Eppen (1979) and Hartman et al. (2000) showed that, in total, it is always cheaper for all outlets to participate in a single, centrally managed inventory system modeled as above, due to the inequality

$$c_R(U) = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij}} = \sqrt{\sum_{i=1}^{n} \sigma_i^2 + 2 \sum_{j>i} \sigma_i \sigma_j \rho_{ij}} \leq \sum_{i=1}^{n} \sigma_i = \sum_{i=1}^{n} c\big(\{i\}\big).$$

The above inequality intuitively suggests that low or negative correlations yield a lower cost to the coalition (when the individual distributions of demand at the outlets, i.e., μ and σ, are fixed). Said differently, certain correlation structures yield lower centralized inventory costs than others. A natural idea is thus: we could lower the coalition's overall expected cost if it were possible to adjust R, while maintaining demands and variances at each outlet.

Is it possible to adjust correlations in such a manner, however? Hartman and Dror (2003) discuss several real-life examples in which companies understand the benefit of—and implement strategies for—reducing these cross correlations (see also the introduction and papers by Gerchak and He 2003; Benjaafar et al. 2005). An assumption of this paper is that it is indeed possible to modify correlations in some interesting settings, and in Sect. 3, we present a correlation matrix R^* that minimizes $c_R(U)$. Some additional comments on modifying correlations are given in Sect. 2.3.

2.2 Allocating centralized inventory costs

Once the correlations between the outlets have been fixed, it is important to allocate the total cost $c_R(U)$ to all outlets in a "fair" manner. For example, one possible allocation would be to charge the entire cost $c_R(U)$ to outlet 1 and zero cost to outlets

$2, \dots, n$. While outlets $2, \dots, n$ would see this allocation as beneficial to them, outlet 1 might judge this to be unfair. If this occurs, outlet 1 would be likely to leave the coalition.

2.2.1 The meaning of "fair"

In order to define what a fair allocation is, we follow Hartman and Dror (2003) and interpret the centralization as a cooperative game, which is defined formally as follows: a *cooperative n-person game on U*, in characteristic function form, is an ordered pair $(U; f)$, where $f : 2^U \to \Re$ is a real-valued set-function on the collection 2^U of all subsets of U such that $f(\emptyset) = 0$. In our inventory model, we may consider $f = c_R$, for example. The corresponding game is referred to as the *cost game* with respect to f.

We also introduce a bit of notation. For any subset $S \subseteq U$ and vector $q \in \Re^n$, we use the notation $q(S)$ to define the sum of the components in q corresponding to the members in S, i.e., $q(S) = \sum_{i \in S} q_i$.

A *cost allocation* is a vector $a = (a_1, \dots, a_n)$ that assigns to each outlet its portion of the cost $f(U)$. Note that a may have negative entries (also referred to as *preimputation*, see Peleg and Sudhölter 2003), unless specifically constrained to be nonnegative. We consider a cost allocation a to be fair with respect to f if it satisfies the following conditions:

1. *Efficiency* (all of the costs are distributed):

$$a(U) = f(U);$$

2. *Stability* (no subset of outlets has an incentive to leave the coalition because it feels it is subsidizing the remaining outlets by paying too much of the cost):

$$a(S) \le f(S) \quad \forall S \subseteq U.$$

Conditions 1 and 2 together define membership in the *core*. In general, it is possible that there exists no fair cost allocation with respect to f.

In this paper, we focus our attention on costs, but a "dual" point of view is to examine the total benefits (or savings) received by a coalition S of outlets when it centrally manages its inventory. This benefit is measured as

$$g(S) = \sum_{i \in S} f\big(\{i\}\big) - f(S).$$

It is also important to distribute the benefits $g(U)$ fairly. A *benefit allocation* is a vector $x = (x_1, \dots, x_n)$ whose coordinates are the benefits received by the members of U, and notions of efficiency and stability for benefits are easily defined: $x(U) = g(U)$ and $x(S) \ge g(S)$ for all $S \subseteq U$. The consideration of benefits leads to an additional notion of fairness:

3. *Justifiability* (at the level of singletons, costs and benefits are compatible):

$$x_i = f\big(\{i\}\big) - a_i \quad \forall i \in U.$$

Because we focus on costs in this paper, given a fair (i.e., efficient and stable) cost allocation a, we define an automatic benefit allocation x by $x_i = f(\{i\}) - a_i$, which satisfies justifiability by definition. Moreover, because the core condition for singleton subsets implies $a_i \le f(\{i\})$, it is not difficult to see that x is also efficient and stable relative to benefits. Approaches for calculating x, which are more sophisticated than ours, may be possible. For example, there are various applications in which one can calculate a and x independently and still achieve justifiability (Hartman and Dror 1996). From this point on, we do not mention benefits explicitly.

There may be multiple cost allocations a that satisfy the above fairness conditions. Schmeidler (1969) described how to compute a specific allocation (either nonnegative or unrestricted) that "makes the least well-off coalition as well-off as possible" (Young 1985a); this allocation is called the *nucleolus*. Following Schmeidler, we focus on the nucleolus as the preferred cost allocation scheme even though a number of other viable allocations exist. We discuss the relationship between the nucleolus and other allocations below.

The nucleolus is calculated by solving a sequence of n successively refined linear programs (LPs), each having roughly 2^n constraints. In the Appendix, we provide a full description of this procedure. In the case of the cost game with respect to f, the initial LP that one must solve is

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & f(S) - a(S) \ge \varepsilon \quad \forall \emptyset \ne S \subsetneq U, \\ & a(U) = f(U), \\ & [a \ge 0]. \end{aligned} \tag{1}$$

The brackets around the constraint $a \ge 0$ indicate that it can be enforced if desired. Note that any feasible solution (a, ε) to this LP automatically satisfies the efficiency condition. Moreover, the core condition is met if and only if $\varepsilon \ge 0$. In this sense, solving (1) can be interpreted as finding an allocation, which is as fair as possible, i.e., an allocation for which ε is as large as possible (even if that ε turns out to be negative).

Besides the nucleolus, there are a host of competing cost allocation methods and choosing among them is somewhat subjective. For instance, in Hartman and Dror (1996) seven different allocation methods are compared in one specific inventory centralization instance. One must ask what are the fundamental properties that an allocation should possess in a given cooperative game setting.

The Shapley value (Shapley 1953) is a common allocation method that assigns a unique allocation to every game. Roughly, this allocation gives each player his average marginal contribution in the game. It is also the unique allocation that satisfies the *dummy axiom*. (A player is a *dummy* if he contributes nothing to any coalition, and the dummy axiom states that a dummy's allocation is zero.) The Shapley value also satisfies the *additivity axiom*. (A cost allocation method ϕ is *additive* if for any joint cost functions c and c' on U, $\phi(c + c') = \phi(c) + \phi(c')$, where $c + c'$ is defined by $(c + c')(S) = c(S) + c'(S)$ for all $S \subseteq U$.) Finally, the Shapley value is *monotonic*. (An allocation method ϕ is *monotonic* if an increase in the cost of a particular coalition implies, ceteris paribus, no decrease in the cost allocation to any members of that coalition.) These three properties are attractive features of the Shapley value.

However, the Shapley value fails to be a core allocation in nonconcave games. Unfortunately, the newsvendor centralization of this paper is not concave; Hartman (1994) provides a simple three player example for this fact. In computational experiments for a nonconcave inventory consolidation ordering game described in Dror et al. (2008), the Shapley value fails to be in the core in about 30% of the cases for which the core is nonempty.

In contrast, the nucleolus is always in the core when the core is nonempty, and even when the core is empty, the nucleolus assigns a cost that cannot be improved by any coalition, thus guaranteeing an elementary condition for stability—an important consideration in a supply chain. The core property of the nucleolus, its "geometric" centrality, and its minimization of the worst impact to any coalition make the nucleolus a fair and defendable cost allocation in a newsvendor centralization situation such as ours. Finally, the nucleolus is non-monotonic. However, Young (1985b) proved that monotonicity is fundamentally incompatible with staying in core by showing that for $|U| \geq 5$ there exists no monotonic core allocation method.

As presented above, the fair allocation of costs clearly depends on the function f being used to measure actual costs for subsets $S \subseteq U$. In this paper, we consider the form c_{R^*}, where R^* is assumed to be an optimal correlation matrix. This function possesses special properties as described in Sect. 4.

2.3 Comments on assumptions of the model

The assumptions of the model explained above are certainly an approximation of real-life. In this paper, we attempt to achieve assumptions that are as realistic as possible, while enabling analysis that sheds light on universal properties of inventory centralization.

The assumption of normal demands, which matches the approach of Eppen (1979) exactly, is key to our entire approach because it enables analytical formulas for inventory costs. It seems difficult to obtain concrete results under other distributions; further research would certainly be interesting.

The assumption of identical holding and penalty costs across all subsets of outlets is also key to our model (and Eppen's) because it allows these costs to be "scaled out." Further, Hartman and Dror have shown in an unpublished note that fair cost allocations do not exist under general holding and penalty costs.

The ability to modify correlations to any desired levels, while maintaining the individual demands at the outlets, is a basic assumption of the results in Sects. 3–4. As mentioned in the introduction, there are many real-life examples in which companies try to manipulate correlations. Even still, modifying correlations precisely is not something that is common managerial expertise. We consider this assumption to be analogous to arguments based on perfect information (PI). Just as arguments with PI allow one to assess the value of PI, our assumption allows us to gauge the value of modifying correlations. In this sense, we can establish fundamental properties of inventory centralization, even if the exact assumptions are realistically hard to achieve.

3 Adjusting the correlations

We assume that μ_i and σ_i are fixed for all $i \in U$. From Sect. 2, our optimization problem is to find R that minimizes $c_R(U)$, or equivalently R that minimizes

$$c_R(U)^2 = \sum_{i=1}^{n} \sigma_i^2 + 2\sum_{j>i} \sigma_i \sigma_j \rho_{ij} = \sigma^T R \sigma.$$

Since the set of correlation matrices can be described as all symmetric, positive semidefinite matrices with ones on the diagonal, our optimization is

$$\begin{aligned} \min \quad & \sigma^T R \sigma \\ \text{s.t.} \quad & \operatorname{diag}(R) = e, \\ & R \succeq 0, \end{aligned} \tag{2}$$

where e is the all-ones vector and the constraint $R \succeq 0$ indicates that R is symmetric and positive semidefinite. As the objective and diagonal constraints are linear in R, (2) is a (linear) semidefinite program (SDP), which is a type of convex programming problem that can be solved up to any desired accuracy in polynomial-time using interior-point methods (Wolkowicz et al. 2000). However, we will establish below an easily computable, closed-form solution of (2), which obviates the need for iterative algorithms in this case. The key idea, echoed in Theorem 3.1, Proposition 3.1, and Theorem 3.2 below, is to lower the overall variance $\sigma^T R \sigma$ of the coalition using antithetic random variables. The multinormal structure of the demand distribution is critical for this.

In the context of inventory centralization, the SDP (2) was introduced by Hartman and Dror (2003), but SDPs having the same constraint structure have been studied for some time. In particular, the set of correlation matrices serves as the basis for relaxations of the maximum-cut problem (Goemans and Williamson 1995), which refers to the decomposition of a graph by deletion of edges of maximum weight. In addition, optimization over correlations matrices has been investigated in the statistics literature on minimum trace factor analysis. In fact, based on the specific structure of the objective function of (2), Shapiro (1982) has provided a partial classification of the optimal solutions of (2). To state the theorem, we make the following assumption without loss of generality:

Assumption *The components of σ are sorted in nonincreasing order.*

Then the theorem of Shapiro (1982) is as follows:

Theorem 3.1 *Let $n \geq 2$, and let $v \in \Re^n$ be the vector having all -1's except for a 1 in the first position. It holds that*:

(a) *If $v^T\sigma > 0$, then $R = vv^T$ is the unique optimal solution of* (2) *with optimal value* $(v^T\sigma)^2$.
(b) *If $v^T\sigma = 0$, then $R = vv^T$ is an optimal solution of* (2) *with optimal value* 0.

(c) *If* $v^T\sigma < 0$, *then the optimal value of* (2) *is* 0, *so that all optimal solutions satisfy the equation* $R\sigma = 0$.

Theorem 3.1 provides valuable information about (2)—and does so practically, since one can easily check the sign of $v^T\sigma$. It is important to point out, however, that Theorem 3.1 does not specifically provide an optimal solution when $v^T\sigma < 0$, i.e., we still must calculate an optimal solution on our own. In the following subsection, we examine simple ways to do so.

Note also that, among the three cases considered by Theorem 3.1, the case $v^T\sigma < 0$ is likely to occur in many real-world situations where all outlets are of roughly the same size and characteristics, i.e., where no single outlet dominates the others. In this sense, $v^T\sigma < 0$ is a highly interesting case, warranting further consideration.

3.1 Calculating an optimal solution when $v^T\sigma < 0$

In this subsection, we assume $v^T\sigma < 0$, where v is as in Theorem 3.1; note that $v^T\sigma < 0$ implies $n \geq 3$. By Theorem 3.1, the optimal value of the inventory centralization is 0. We wish to calculate an optimal correlation matrix R satisfying $\sigma^T R\sigma = 0$ or, equivalently, $R\sigma = 0$.

Our analysis will be based on the fundamental fact that any symmetric positive semidefinite R can be factored as $R = VV^T$ for some $V \in \Re^{n\times n}$. Note that the factorization is not unique in general. Thus, our search for a correlation matrix R can be cast as a search for V with unit-length rows such that $\sigma^T VV^T\sigma = \|\sigma^T V\|^2 = 0$, where $\|\cdot\|$ is the standard vector Euclidean norm. Equivalently, we search for V satisfying $\sigma^T V = 0$.

The factor V can in fact be taken to have rank(R) columns. However, we have no prior knowledge of this rank except for the following: rank$(R) \leq n-1$ because σ is a nonzero vector in the null space of R. So without loss of generality, we can restrict our search to $V \in \Re^{n\times(n-1)}$. In the results below, we will in fact demonstrate an optimal R such that rank$(R) \leq 2$.

3.1.1 The case for $n = 3$

We consider first the case when $n = 3$. Given $\sigma = (\sigma_1, \sigma_2, \sigma_3)^T$, define

$$R^* = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix} \quad \text{and} \quad V^* = \begin{pmatrix} 1 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} \\ \rho_{13} & -\sqrt{1-\rho_{13}^2} \end{pmatrix}, \tag{3}$$

where

$$\rho_{12} = \left(\sigma_3^2 - \sigma_1^2 - \sigma_2^2\right)/(2\sigma_1\sigma_2),$$
$$\rho_{13} = \left(\sigma_2^2 - \sigma_1^2 - \sigma_3^2\right)/(2\sigma_1\sigma_3),$$
$$\rho_{23} = \left(\sigma_1^2 - \sigma_2^2 - \sigma_3^2\right)/(2\sigma_2\sigma_3).$$

We have the following proposition:

Proposition 3.1 *Suppose* $\sigma \in \Re^3$ *satisfies* $\sigma > 0$ *and* $v^T\sigma < 0$, *and define* R^* *and* V^* *by* (3). *Then* $R^* = V^*(V^*)^T$, *and* R^* *is the unique optimal solution of* (2). *Specifically*, $\sigma^T R^* \sigma = 0$ *and* $\sigma^T V^* = 0$.

Proof We remark that the assumption $v^T\sigma < 0$ implies any optimal solution to (2) must have rank equal to 2, which is consistent with the proposed structure for V^* (i.e., 2 columns).

We first show that V^* is well defined, i.e., that $|\rho_{12}| \le 1$ and $|\rho_{13}| \le 1$. In particular, we show $|\rho_{12}| < 1$ and $|\rho_{13}| < 1$. Since ρ_{12} and ρ_{13} are defined similarly, we give the argument for ρ_{12} only. If the numerator of ρ_{12} is nonnegative, then it suffices to show $\rho_{12} < 1$, which is equivalent to

$$\sigma_3^2 - \sigma_1^2 - \sigma_2^2 < 2\sigma_1\sigma_2 \quad \Longleftrightarrow \quad \sigma_3^2 < (\sigma_1 + \sigma_2)^2 \quad \Longleftrightarrow \quad \sigma_3 < \sigma_1 + \sigma_2,$$

which is true because σ is sorted in descending order. On the other hand, if the numerator is negative, then we require $\rho_{12} > -1$, which is equivalent to

$$\begin{aligned} \sigma_3^2 - \sigma_1^2 - \sigma_2^2 > -2\sigma_1\sigma_2 \quad &\Longleftrightarrow \quad \sigma_3^2 > (\sigma_1 - \sigma_2)^2 \\ &\Longleftrightarrow \quad \sigma_3 > \max\{\sigma_1 - \sigma_2, \sigma_2 - \sigma_1\}, \end{aligned}$$

which is also true because $v^T\sigma < 0$ and because σ is sorted. It follows that $|\rho_{12}| < 1$.

We next argue $R^* = V^*(V^*)^T$. It is clear that $\operatorname{diag}(V^*(V^*)^T) = e$, $(V^*(V^*)^T)_{12} = \rho_{12}$, and $(V^*(V^*)^T)_{13} = \rho_{13}$, and so it remains to show

$$\rho_{12}\rho_{13} - \sqrt{1 - \rho_{12}^2}\sqrt{1 - \rho_{13}^2} = \rho_{23}. \tag{4}$$

We first remark that one can readily show $\sigma_2^2(1 - \rho_{12}^2) = \sigma_3^2(1 - \rho_{13}^2)$, which implies the left-hand side of (4) equals $\rho_{12}\rho_{13} - (\sigma_2/\sigma_3)(1 - \rho_{12}^2)$. This expression in turn simplifies to ρ_{23}.

To complete the proof of the proposition, we demonstrate that any $V \in \Re^{3\times 2}$ having unit-length rows and satisfying $\sigma^T V = 0$ must yield $R^* = VV^T$. We note that, for all orthogonal $Q \in \Re^{2\times 2}$: (i) $VV^T = (VQ)(VQ)^T$; and (ii) the equation $\sigma^T V = 0$ holds if and only if $\sigma^T VQ = 0$. As a result, we may assume without loss of generality that a rotation has been applied to the rows of V so that $V_{1\cdot} = (1, 0)$. We write

$$V = \begin{pmatrix} 1 & 0 \\ a & b \\ c & d \end{pmatrix}$$

and require

$$a^2 + b^2 = 1, \qquad c^2 + d^2 = 1, \qquad \sigma_1 + \sigma_2 a + \sigma_3 c = 0, \qquad \sigma_2 b + \sigma_3 d = 0.$$

These equations imply

$$a^2 + (\sigma_3/\sigma_2)^2 d^2 = 1, \qquad (1/\sigma_3)^2(\sigma_1 + \sigma_2 a)^2 + d^2 = 1,$$

which in turn imply $a = \rho_{12}$ after substituting for d^2. A similar argument yields $c = \rho_{13}$. Since the equation $\sigma_2 b + \sigma_3 d = 0$ implies that b and d have opposite signs, we have

$$(b, d) = \pm\big(\sqrt{1-\rho_{12}^2}, -\sqrt{1-\rho_{13}^2}\big).$$

In either case, the product VV^T equals R^*. □

Proposition 3.1 provides the optimal solution for the case when $n = 3$, and although this result may seem specialized, we show in the next subsection that it actually allows us to solve (2) easily for arbitrary n.

3.1.2 The case for arbitrary $n \geq 3$

Let $\mathcal{S}$ be a bi- or tri-partition of $\{1, \dots, n\}$; we write $\mathcal{S} = \{S_1, S_2\}$ or $\mathcal{S} = \{S_1, S_2, S_3\}$ as appropriate. Define

$$\bar{\sigma} \in \Re^{|\mathcal{S}|}, \qquad \bar{\sigma}_j = \sigma(S_j) = \sum_{k \in S_j} \sigma_k, \quad 1 \leq j \leq |\mathcal{S}|. \tag{5}$$

We say that $\mathcal{S}$ is *balanced* if

$$\bar{\sigma}_1 = \bar{\sigma}_2 \quad \text{when } |\mathcal{S}| = 2$$

or if

$$\bar{\sigma}_1 < \bar{\sigma}_2 + \bar{\sigma}_3, \qquad \bar{\sigma}_2 < \bar{\sigma}_1 + \bar{\sigma}_3, \qquad \bar{\sigma}_3 < \bar{\sigma}_1 + \bar{\sigma}_2 \quad \text{when } |\mathcal{S}| = 3. \tag{6}$$

Intuitively, a balanced $\mathcal{S}$ divides the outlets into groups, where the total amount of standard deviation of demand in each group is equal—or not excessively high. The following result, the proof of which is constructive, establishes that σ admits a balanced partition.

Proposition 3.2 *Let $n \geq 3$, and suppose $v^T\sigma < 0$. Then there exists a balanced bi-partition $\mathcal{S} = \{S_1, S_2\}$ or a balanced tri-partition $\mathcal{S} = \{S_1, S_2, S_3\}$ of $\{1, \dots, n\}$ with respect to σ.*

Proof Because σ is sorted in nonincreasing order, it is trivial to determine whether there exists k such that $\sigma_1 + \cdots + \sigma_k = \sigma_{k+1} + \cdots + \sigma_n$. If so, then define $\mathcal{S} = \{S_1, S_2\}$, where $S_1 = \{1, \dots, k\}$ and $S_2 = \{k+1, \dots, n\}$; $\mathcal{S}$ is the required bi-partition.

Otherwise, let k be the index such that

$$\sigma_1 + \cdots + \sigma_{k-1} < \sigma_k + \sigma_{k+1} + \cdots + \sigma_n,$$
$$\sigma_1 + \cdots + \sigma_{k-1} + \sigma_k > \sigma_{k+1} + \cdots + \sigma_n.$$

We certainly have $k \geq 2$ because $v^T\sigma < 0$; we also have $k \leq n-1$ because the components of σ are sorted and because $n \geq 3$. Now define $\mathcal{S}$ by

$$S_1 = \{1, \dots, k-1\}, \qquad S_2 = \{k\}, \qquad S_3 = \{k+1, \dots, n\},$$

and also define $\bar{\sigma}$ according to (5). We claim that $\mathcal{S}$ is balanced. The first inequality above shows $\bar{\sigma}_1 < \bar{\sigma}_2 + \bar{\sigma}_3$, while the second inequality shows $\bar{\sigma}_3 < \bar{\sigma}_1 + \bar{\sigma}_2$. Finally, we have

$$\bar{\sigma}_2 = \sigma_k \le \sigma_1 < (\sigma_1 + \cdots + \sigma_{k-1}) + (\sigma_{k+1} + \cdots + \sigma_n) = \bar{\sigma}_1 + \bar{\sigma}_3,$$

which follows because σ is sorted and because $n \ge 3$. So $\mathcal{S}$ is the desired tri-partition. □

We remark that, in the case of a balanced tri-partition, we may assume without loss of generality that $\mathcal{S}$ has been defined so that the components of $\bar{\sigma}$ are sorted in nonincreasing order. Moreover, if we let $\bar{v} = (1, -1, -1)^T$ be as in Theorem 3.1 relative to $\bar{\sigma}$, then the three conditions (6) are equivalent to the single inequality $\bar{v}^T \bar{\sigma} < 0$.

Using Propositions 3.1 and 3.2, the following theorem shows how to construct an optimal solution of (2) easily.

Theorem 3.2 *Let $n \ge 3$, and suppose that $\sigma \in \Re^n$ satisfies $\sigma > 0$ and $v^T \sigma < 0$. Let $\mathcal{S}$ be any balanced bi- or tri-partition relative to σ, and define $\bar{\sigma}$ by* (5). *Define $\bar{V}^* \in \Re^{|\mathcal{S}| \times (|\mathcal{S}|-1)}$ as follows*:

(i) *If $|\mathcal{S}| = 2$, then $\bar{V}^* = (1, -1)^T$.*
(ii) *If $|\mathcal{S}| = 3$, then $\bar{V}^*$ is as defined by* (3) *with respect to $\bar{\sigma}$.*

In addition, define $V^* \in \Re^{n \times (|\mathcal{S}|-1)}$ row-by-row as

$$V^*_{i\cdot} = \bar{V}^*_{j\cdot} \quad \text{if } i \in S_j \ \big(1 \le i \le n,\ 1 \le j \le |\mathcal{S}|\big), \tag{7}$$

and $R^* = V^*(V^*)^T$. Then R^* is a rank-$(|\mathcal{S}|-1)$ optimal solution of (2). *In particular, $\sigma^T R^* \sigma = 0$ and $\sigma^T V^* = 0$.*

Proof By construction, each row of V^* has unit norm so that R^* is feasible for (2). It remains to show that R^* is optimal, which is implied by the following equation:

$$\begin{aligned} \sigma^T V^* &= \sum_{i=1}^n \sigma_i V^*_{i\cdot} = \sum_{j=1}^{|\mathcal{S}|} \sum_{i \in S_j} \sigma_i V^*_{i\cdot} = \sum_{j=1}^{|\mathcal{S}|} \sum_{i \in S_j} \sigma_i \bar{V}^*_{j\cdot} \\ &= \sum_{j=1}^{|\mathcal{S}|} \sigma(S_j) \bar{V}^*_{j\cdot} = \sum_{j=1}^{|\mathcal{S}|} \bar{\sigma}_j \bar{V}^*_{j\cdot} = \bar{\sigma}^T \bar{V}^* = 0. \end{aligned}$$

□

We remark that Theorem 3.2 provides only a single optimal solution R^* of (2) and that R^* is dependent on the partition $\mathcal{S}$. In general, (2) may have multiple optimal solutions. However, R^* is "minimal" in the sense that its rank, which equals 1 or 2, is the smallest of all optimal solutions.

On a related note, we mention that, if (2) is solved via an off-the-shelf SDP solver—rather than via the closed-form solution provided by Theorem 3.2—then in general one would receive a high-rank optimal solution since such solvers work with full-rank interior points. Indeed, such a high-rank solution may be a convex combination of several low-rank solutions. In this sense, solving (2) via a standard SDP solver does not reveal the structure of Theorem 3.2.

Theorem 3.2 can also be compared to Pataki (1998), which guarantees the existence of an optimal solution of (2) having rank approximately $\sqrt{2n}$. Pataki's rank result holds for any linear objective in (2), whereas our rank result uses the structure of the specific objective of (2). Hence, our rank bound is much tighter than Pataki's in this context.

3.2 Examples

To illustrate the results of Theorems 3.1 and 3.2, we consider the following three examples with $U = \{1, \ldots, 6\}$, which differ only in the standard deviation of outlet 1:

$$\sigma^1 = (30, 5, 4, 3, 2, 1),$$
$$\sigma^2 = (15, 5, 4, 3, 2, 1),$$
$$\sigma^3 = (6, 5, 4, 3, 2, 1).$$

To fix concepts, let R be the non-optimized correlation matrix

$$\begin{pmatrix} 1.0 & -0.2 & 0.0 & 0.4 & -0.2 & -0.6 \\ -0.2 & 1.0 & -0.4 & 0.2 & 0.0 & 0.3 \\ 0.0 & -0.4 & 1.0 & -0.4 & 0.6 & -0.2 \\ 0.4 & 0.2 & -0.4 & 1.0 & -0.2 & -0.5 \\ -0.2 & 0.0 & 0.6 & -0.2 & 1.0 & 0.0 \\ -0.6 & 0.3 & -0.2 & -0.5 & 0.0 & 1.0 \end{pmatrix}.$$

Then the total centralization cost $c_R(U)$ in each example is

$$\left[(\sigma^1)^T R \sigma^1\right]^{1/2} = 29.8331,$$
$$\left[(\sigma^2)^T R \sigma^2\right]^{1/2} = 15.5563,$$
$$\left[(\sigma^3)^T R \sigma^3\right]^{1/2} = 8.2098.$$

Notice also that the cost $c_R(S)$ to the subset coalition $S = \{2, \ldots, 6\}$ is the same in each example, namely 6.4031.

Now we consider the optimal R^* in each example. Let $v = (1, -1, -1, -1, -1, -1)^T$, and note that $v^T \sigma^1 > 0$, $v^T \sigma^2 = 0$, and $v^T \sigma^3 < 0$. Then, according to Theorem 3.1, we have the following table of (partial) results:

Example	$v^T \sigma$	R^*	$c_{R^*}(U)$
σ^1	positive	vv^T	$v^T \sigma^1$
σ^2	zero	vv^T	0
σ^3	negative	(n/a)	0

It is worthwhile to examine the explicit form of R^* in the case of σ^1 and σ^2:

$$R^* = \begin{pmatrix} 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

So outlet 1 is perfectly negatively correlated with outlets 2 through 6, and outlets 2 through 6 are perfectly positively correlated amongst themselves. Note also that Theorem 3.1 does not provide the optimal solution for the case of σ^3, even though it does provide the minimum cost, which is the gap that Theorem 3.2 addresses.

To calculate R^* for the σ^3 case, Theorem 3.2 requires that we first construct a balanced bi- or tri-partition of $\{1, \dots, 6\}$ with respect to σ^3. For example,

$$\mathcal{S} = \{S_1, S_2, S_3\} = \big\{\{3, 4, 5, 6\}, \{1\}, \{2\}\big\}$$

is a balanced tri-partition with

$$\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2, \bar{\sigma}_3)^T = \big(\sigma(S_1), \sigma(S_2), \sigma(S_3)\big)^T = (10, 6, 5)^T.$$

Then, according to (3),

$$\bar{V}^* = \begin{pmatrix} 1.0000 & 0 \\ -0.9250 & 0.3800 \\ -0.8900 & -0.4560 \end{pmatrix}$$

yields the optimal correlation matrix $\bar{R}^*$ with respect to $\bar{\sigma}$ via the formula

$$\bar{R}^* = \bar{V}^*\big(\bar{V}^*\big)^T = \begin{pmatrix} 1.0000 & -0.9250 & -0.8900 \\ -0.9250 & 1.0000 & 0.6500 \\ -0.8900 & 0.6500 & 1.0000 \end{pmatrix}.$$

With this information in hand, Theorem 3.2 "pulls back" to determine an optimal solution with respect to σ^3:

$$V^* = \begin{pmatrix} -0.9250 & 0.3800 \\ -0.8900 & -0.4560 \\ 1.0000 & 0 \\ 1.0000 & 0 \\ 1.0000 & 0 \\ 1.0000 & 0 \end{pmatrix}$$

and

$$R^* = V^*(V^*)^T$$
$$= \begin{pmatrix} 1.0000 & 0.6500 & -0.9250 & -0.9250 & -0.9250 & -0.9250 \\ 0.6500 & 1.0000 & -0.8900 & -0.8900 & -0.8900 & -0.8900 \\ -0.9250 & -0.8900 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ -0.9250 & -0.8900 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ -0.9250 & -0.8900 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ -0.9250 & -0.8900 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \end{pmatrix}.$$

One can also check that $c_{R^*}(U) = [(\sigma^3)^T R^* \sigma^3]^{1/2}$ equals 0, as predicted.

3.3 Some comments on managerial decision making

As mentioned in Sect. 2.3, we consider the assumption of the ability to modify correlations in an operational setting to be analogous to assumptions of perfect information in stochastic situations. In particular, although it may not always be possible for managers to modify correlations precisely, assuming this ability allows us to assess the potential value of being able to do so. We believe the results of the previous subsections are properly viewed through this lens.

Indeed, Theorems 3.1 and 3.2 show the substantial potential value of correlation modification: inventory costs can be minimized to $\max\{0, \sigma_1 - (\sigma_2 + \cdots + \sigma_n)\}$ (assuming that the components of σ are sorted in nonincreasing order). Theorems 3.1 and 3.2 also give insights on how managers might work toward the optimal cost in an operational setting, as we explain next.

First, the optimal correlation matrix R^* can serve as a desired target for a manager. Although it may be unrealistic to expect to achieve R^* exactly, R^* does present a tangible example of improved correlations. Moreover, if R is the current correlation structure, the manager can be sure that any movement in the direction $R^* - R$ will reduce expected inventory costs.

Second, in each of the cases considered ($v^T\sigma$ positive, zero, or negative), the optimal solution R^* provided is always of very low-rank (either rank 1 or rank 2 according to Theorems 3.1 and 3.2). This structure on R^* is somewhat non-intuitive. Why should the optimal solution have such a property, particularly independent of the number of outlets? We unfortunately do not have a satisfactory a priori explanation, but it is interesting to understand the implications of this result.

The low-rank structure of R^* shows that optimal centralization is achieved when the outlets are partitioned into groups in accordance with the following guideline: for any particular group, the sum of the standard deviations of the outlets in the group should not be too large compared to the same measure for other groups. Said differently, one group's cost should not be too large compared to that of other groups. When $v^T\sigma \geq 0$, the groups are $\{1\}$ and $\{2, \ldots, n\}$; when $v^T\sigma < 0$, the groups correspond to the balanced bi- or tri-partition of Theorem 3.2. Furthermore, the low-rank structure of R^* demonstrates that, in a sense, the correlations should be with respect to groups only. All outlets within a group should be perfectly (positively) correlated, and for each pair of groups, there should be a single correlation, common to all pairs of outlets between the groups. The examples of the previous subsection illustrate these points.

For the manager trying to reduce his/her centralization costs, this gives a conceptually easy strategy to follow (even if managing correlations precisely is difficult). First, partition outlets into a few groups so that costs (standard deviations) are balanced. Then attempt to correlate the groups so as to minimize costs, using techniques discussed in the introduction, for example. The low-rank matrix R^* can serve as a guide. In particular, it is not necessary to micromanage all $n(n-1)/2$ pairs of correlations.

Finally, a manager may be interested in alternate optimal solutions R^*, which could be evaluated relative to some additional “soft constraints” (such as closeness to the current correlation structure R). Using alternate balanced partitions, which are easy to generate, these alternate solutions can be calculated in a straightforward manner.

4 Fair allocations under optimized correlations

In this subsection, we assume that we have determined R^* that minimizes the total centralization cost $c_R(U)$, without affecting the demand distribution at each outlet (see Sect. 3). With correlations fixed as R^*, we now turn our attention to allocating the costs and benefits in a fair fashion (see Sect. 2).

By Theorem 3.1, we know that the optimal cost is positive if $v^T\sigma > 0$ and 0 otherwise. In the second case, $a = 0$ is an obvious fair cost allocation. When only nonnegative allocations are considered, $a = 0$ is in fact the only feasible allocation and so is also the cost nucleolus.

For the case when $c_{R^*}(U) = v^T\sigma > 0$ and $R^* = vv^T$ is the unique optimal solution, the nucleolus can be computed in a closed form, as shown in Theorem 4.1 below. We remark that the theorem considers both nonnegative and unrestricted allocations simultaneously and provides a nonnegative allocation that is the nucleolus in both situations. For the statement of the theorem, note that e_1 is the first coordinate vector.

Theorem 4.1 *If $v^T\sigma > 0$ and, according to Theorem 3.1, $R^* = vv^T$ and $c_{R^*}(U) = v^T\sigma$, then the nucleolus of the cost game is $(v^T\sigma)e_1$.*

The proof of this theorem is given in the Appendix.

This result says that outlet 1 should pay the entire costs of the centralization. One intuitive explanation of the result is as follows. Because outlet 1 has a high standard deviation of demand compared to all other outlets (i.e., $v^T\sigma > 0$), the overall (optimized) cost $c_{R^*}(U)$ is positive; if σ_1 were smaller (specifically if $v^T\sigma$ were nonpositive), then the overall (optimized) cost would be zero. Hence, the positive cost can be attributed to (or “blamed on”) outlet 1, and so in fairness, outlet 1 should absorb all costs. It is important to keep in mind that outlet 1 still benefits from the coalition in that it is paying $\sigma_2 + \cdots + \sigma_n$ less than it would if it were to act alone.

5 Summary

The main purpose of this paper is to examine the cost optimization problem implied by the newsvendor centralization arrangement when individual outlets experience normally distributed demand and their inventory cost parameters are identical. In addition, it is paramount that the cost of the centralized arrangement be fairly assessed to the individual outlets. The cost optimization problem has the form of a semidefinite program of Goemans and Williamson (1995)—optimization over correlation matrices. Tracing back to the statistics literature (Shapiro 1982) allows us to obtain a partial classification of the optimal solutions of our SDP (Theorem 3.1). We complete the classification of the optimal solutions by constructing an optimal correlation matrix for any balanced partition relative to σ (Theorem 3.2).

This leads to the problem of calculating a fair cost allocation, expressed here as the calculation of the nucleolus for the corresponding cooperative game. In principle, it is a tedious process of solving n linear programs with an exponential number of constraints. However, for the optimal solution of the corresponding SDP, we prove that this can be done easily in a closed form (Theorem 4.1).

From the current study, there are some interesting open questions and possible extensions. For example, newsvendor centralization problems with identical cost parameters are known to have a nonempty core, while the correlation optimization problem examined in this paper is restricted to normally distributed demands. For what other interesting demand distributions can analysis like that of the paper be extended? In addition, this paper has provided the solution of the "unconstrained" correlation optimization, i.e., where we are allowed to shift correlations arbitrarily so as to minimize cost. In reality, the ability to alter correlations could be limited in some manner, which would constrain the SDP (2). It would be interesting to investigate such situations, for which a general purpose SDP solver is likely to be required.

Acknowledgements The authors are in debt to Alexander Shapiro for pointing out Theorem 3.1 and to two anonymous referees who provided many helpful comments and suggestions.

Appendix: Proof of Theorem 4.1

We first discuss the procedure of Schmeidler (1969) for determining the nucleolus. Recall that the nucleolus is calculated by solving a sequence of n linear programs (LPs), starting with (1). Generally, once the kth LP has been solved, the $(k+1)$st LP is constructed as follows. Let ε_k be the optimal value of the kth LP, and let $\mathcal{P}^k$ denote the collection of proper subsets $S \notin \mathcal{P}^1 \cup \cdots \cup \mathcal{P}^{k-1}$ of $\{1, \ldots, n\}$ for which the inequality

$$c_{R^*}(S) - a(S) \geq \varepsilon$$

is active in all optimal solutions of the kth LP. Then the $(k+1)$st LP is

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & c_{R^*}(S) - a(S) \geq \varepsilon \quad \forall S \notin \mathcal{P}^1 \cup \cdots \cup \mathcal{P}^k,\ \emptyset \neq S \subsetneq U, \\ & c_{R^*}(S) - a(S) = \varepsilon_\ell \quad \forall S \in \mathcal{P}^\ell,\ \ell = 1, \ldots, k, \\ & a(U) = c_{R^*}(U). \end{aligned} \tag{8}$$

It is proven by Schmeidler (1969) that the nth LP is guaranteed to have a unique optimal solution, which is the nucleolus.

Our next step is to specialize the first LP (1) to the case of Theorem 4.1, i.e., when $R^* = vv^T$ is the unique optimal solution of (2) and the optimal cost $c_{R^*}(U)$ equals $v^T\sigma = \sigma_1 - (\sigma_2 + \cdots + \sigma_n) > 0$. For this, the first important observation is that

$$c_{R^*}(S) = \left|\sum_{i \in S} \sigma_i v_i\right| = \begin{cases} \sigma_1 - \sigma(S \setminus 1) & \text{if } 1 \in S, \\ \sigma(S) & \text{if } 1 \notin S. \end{cases}$$

Defining $\mathcal{S}_1$ to be the collection of all proper subsets S of $\{1, \ldots, n\}$ having $1 \in S$ and $\mathcal{S}_1^c$ to be the collection of S such that $1 \notin S$, (1) becomes

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & \sigma_1 - \sigma(S \setminus 1) - a(S) \geq \varepsilon \quad \forall S \in \mathcal{S}_1, \\ & \sigma(S) - a(S) \geq \varepsilon \quad \forall S \in \mathcal{S}_1^c, \\ & a(U) = v^T\sigma. \end{aligned} \tag{9}$$

Also define $\mathcal{T}_1$ and $\mathcal{T}_1^c$ similarly to $\mathcal{S}_1$ and $\mathcal{S}_1^c$ except with respect to the ground set $\{1, \ldots, n-1\}$. We consider the problem

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & (\sigma_1 - \sigma_n) - \sigma(S \setminus 1) - a(S) \geq \varepsilon \quad \forall S \in \mathcal{T}_1, \\ & \sigma(S) - a(S) \geq \varepsilon \quad \forall S \in \mathcal{T}_1^c, \\ & a(U \setminus n) = v^T\sigma, \end{aligned} \tag{10}$$

which is the first-stage LP for calculating the nucleolus of the reduced-dimension cost game given by

$$\bar{\sigma} = (\sigma_1 - \sigma_n, \sigma_2, \ldots, \sigma_{n-1})^T \in \Re^{n-1} \tag{11}$$

with the same cost structure. The proof of Theorem 4.1 hinges on establishing the following result, whose proof is given in Sect. 6.1.

Lemma 6.1 *Let σ be given as in Theorem 4.1. During the process for calculating the nucleolus a of the cost game with respect to σ, the second-stage LP ensures $a_n = 0$. Moreover, the second-stage LP is equivalent to (10).*

The equivalence just stated in Lemma 6.1 has the following precise meaning: the feasible sets of (10) and the second-stage LP are in bijective correspondence via the map $(a_1, \ldots, a_{n-1}) \leftrightarrow (a_1, \ldots, a_{n-1}, 0)$.

With Lemma 6.1, Theorem 4.1 is easy to prove:

Proof Let a be the nucleolus. Lemma 6.1 establishes that $a_n = 0$. Moreover, the second-stage LP is equivalent to (10) by Lemma 6.1, where (10) is itself the first-stage nucleolus LP for $\bar{\sigma}$ given by (11) with the same cost structure. Further, $v^T\sigma > 0$ implies that $\bar{\sigma}$ has the analogous property $\bar{\sigma}_1 > \bar{\sigma}_2 + \cdots + \bar{\sigma}_{n-1}$. It then follows by induction that $0 = a_{n-1} = \cdots = a_2$, which proves the theorem. □

6.1 Proof of Lemma 6.1

We will prove Lemma 6.1 in several steps, each of which helps characterize the second-stage LP of the nucleolus calculation. To simplify the presentation, we consider only the case that the allocation a is unrestricted (the proofs for $a \geq 0$ are exactly the same), and we also assume $\sigma_{n-1} > \sigma_n$ (it is not hard to handle the case when the tail of σ is constant).

First, we show that $((v^T\sigma)e_1, \sigma_n)$ is an optimal solution of the first-stage LP.

Lemma 6.2 *The vector $(a, \varepsilon) = ((v^T\sigma)e_1, \sigma_n)$ is an optimal solution of the first-stage LP (9).*

Proof We first show that the proposed vector (a, ε) is feasible. Clearly $a(U) = v^T\sigma$. For convenience, let $[n]$ denote the set $\{1, \ldots, n\}$. Then, for $S \in \mathcal{S}_1$, we have

$$\begin{aligned}\sigma_1 - \sigma(S \setminus 1) - a(S) &= \sigma_1 - \sigma(S \setminus 1) - v^T\sigma = \sigma_1 - \sigma(S \setminus 1) - \big(\sigma_1 - \sigma\big([n] \setminus 1\big)\big)\\ &= \sigma\big([n] \setminus 1\big) - \sigma(S \setminus 1) = \sigma\big([n] \setminus S\big) \geq \sigma_n,\end{aligned}$$

where the last inequality follows because $[n] \setminus S$ is nonempty and σ_n is the smallest component in σ. Next, for $S \in \mathcal{S}_1^c$, we have

$$\sigma(S) - a(S) = \sigma(S) \geq \sigma_n,$$

which follows because S is nonempty and σ_n is smallest. Overall, we see that the proposed (a, ε) is feasible.

We can also show that any feasible (a, ε) has $\varepsilon \leq \sigma_n$, which will prove the result. Taking $S = \{n\}$ and considering the constraint $\sigma(S) - a(S) \geq \varepsilon$, we see $\sigma_n - a_n \geq \varepsilon$. In addition, taking $S = \{1, \ldots, n-1\}$ and considering the constraint $\sigma_1 - \sigma(S \setminus 1) - a(S) \geq \varepsilon$, we see that

$$\sigma_1 - (\sigma_2 + \cdots + \sigma_{n-1}) - (a_1 + \cdots + a_{n-1}) \geq \varepsilon.$$

Adding these two inequalities and using the fact that $a(U) = v^T\sigma = \sigma_1 - (\sigma_2 + \cdots + \sigma_n)$, we conclude $2\sigma_n \geq 2\varepsilon$, as desired. □

In accordance with the procedure for calculating the nucleolus, our next step is to determine which inequalities are active in every optimal solution of (9). We do so using a duality argument. Introducing a dual variable $y_S \geq 0$ for each inequality constraint of (9) and a free dual variable λ for the equality constraint, the dual of (9) is

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}_1} [\sigma_1 - \sigma(S \setminus 1)] y_S + \sum_{S \in \mathcal{S}_1^c} \sigma(S) y_S - (v^T \sigma)\lambda \\ \text{s.t.} \quad & \lambda = \sum_{S \in \mathcal{S}_j} y_S, \quad \forall j = 1, \dots, n, \\ & \sum_{\emptyset \neq S \subsetneq U} y_S = 1, \quad y \geq 0, \end{aligned} \tag{12}$$

where, analogous to $\mathcal{S}_1$, for all $j = 2, \dots, n$, $\mathcal{S}_j$ is defined to be the collection of all proper subsets S of $\{1, \dots, n\}$ having $j \in S$.

Lemma 6.3 *Consider the optimal solution* $(a, \varepsilon) = ((v^T \sigma)e_1, \sigma_n)$ *of* (9). *The active inequalities correspond precisely to the sets* $\{1, \dots, n-1\}$ *and* $\{n\}$. *As a result, in every optimal dual solution* (y, λ) *of* (12), *it holds that* $y_S = 0$ *for all other sets* S.

Proof The proof of Lemma 6.2 shows that, with the optimal solution (a, ε), the inequalities of (9) reduce to

$$\begin{aligned} &\sigma([n] \setminus S) \geq \sigma_n \quad \forall S \in \mathcal{S}_1, \\ &\sigma(S) \geq \sigma_n \quad \forall S \in \mathcal{S}_1^c. \end{aligned}$$

From this, it is not difficult to see that equality is attained—and can only be attained—for the sets $\{1, \dots, n-1\}$ and $\{n\}$. Complementary slackness proves the second part of the lemma. □

Next, we show that the same inequalities are active even in the case of multiple optimal solutions.

Proposition 6.1 *Regarding* (9), *the inequalities which are active in every optimal solution correspond precisely to the sets* $\{1, \dots, n-1\}$ *and* $\{n\}$.

Proof We know from Lemma 6.3 that those inequalities *not* corresponding to $\{1, \dots, n-1\}$ or $\{n\}$ are inactive in the specific optimal solution provided by Lemma 6.2. To prove the current proposition, it thus remains to show that the inequalities corresponding to $\{1, \dots, n-1\}$ or $\{n\}$ are active in all optimal solutions of (9).

Consider the dual (12). Lemma 6.3 implies that, at optimality, the dual simplifies to

$$\begin{aligned} \min \quad & [\sigma_1 - \sigma(\{2, \dots, n-1\})] y_{\{1, \dots, n-1\}} + \sigma_n y_{\{n\}} - (v^T \sigma)\lambda \\ \text{s.t.} \quad & \lambda = y_{\{1, \dots, n-1\}}, \\ & \lambda = y_{\{n\}}, \\ & y_{\{1, \dots, n-1\}} + y_{\{n\}} = 1 \quad y \geq 0. \end{aligned}$$

It follows that $\lambda = y_{\{1,\ldots,n-1\}} = y_n = 1/2$ with objective value σ_n. Thus, complementary slackness implies that the inequalities in (9) corresponding to $\{1,\ldots,n-1\}$ and $\{n\}$ are active in all optimal solutions of (9). This proves the desired result. □

Now, with Proposition 6.1 in hand, we can construct the second-stage nucleolus LP:

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & \sigma_1 - \sigma(S \setminus 1) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1 \setminus \big\{\{1,\ldots,n-1\}\big\}, \\ & \sigma(S) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1^c \setminus \big\{\{n\}\big\}, \\ & \sigma_1 - \sigma\big(\{2,\ldots,n-1\}\big) - a\big(\{1,\ldots,n-1\}\big) = \sigma_n, \\ & \sigma\big(\{n\}\big) - a\big(\{n\}\big) = \sigma_n, \\ & a(U) = v^T\sigma. \end{aligned} \tag{13}$$

Note that the fourth constraint implies $a_n = 0$ so that the third is implied by $a(U) = v^T\sigma$. As a result, (13) simplifies to

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & \sigma_1 - \sigma(S \setminus 1) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1 \setminus \big\{\{1,\ldots,n-1\}\big\}, \\ & \sigma(S) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1^c \setminus \big\{\{n\}\big\}, \\ & a(U) = v^T\sigma, \\ & a_n = 0. \end{aligned} \tag{14}$$

It is possible to simplify (14) further. Related to the first constraint, consider $S \in \mathcal{S}_1 \setminus \{\{1,\ldots,n-1\}\}$ such that $n \notin S$. We claim that the constraint for $S \cup n$ implies the constraint for S. This is seen by using $a_n = 0$ to establish

$$\begin{aligned} \sigma_1 - \sigma(S \cup n \setminus 1) - a(S \cup n) \ge \varepsilon \quad & \Longleftrightarrow \quad \sigma_1 - \sigma(S \setminus 1) - a(S) \ge \varepsilon + \sigma_n \\ & \Longrightarrow \quad \sigma_1 - \sigma(S \setminus 1) - a(S) \ge \varepsilon. \end{aligned}$$

Related to the second constraint, consider $S \in \mathcal{S}_1^c \setminus \{\{n\}\}$ such that $n \notin S$. In this case, we claim that the constraint for S implies the constraint for $S \cup n$:

$$\begin{aligned} \sigma(S) - a(S) \ge \varepsilon \quad & \Longrightarrow \quad \sigma(S) - a(S) \ge \varepsilon - \sigma_n \\ & \Longleftrightarrow \quad \sigma(S \cup n) - a(S \cup n) \ge \varepsilon. \end{aligned}$$

As a result, we can remove all inequality constraints of the first type for which $n \notin S$ and all constraints of the second type which have $n \in S$ so that (14) becomes

$$\begin{aligned} \max \quad & \varepsilon \\ \text{s.t.} \quad & \sigma_1 - \sigma(S \setminus 1) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1 \setminus \big\{\{1,\ldots,n-1\}\big\} \text{ with } n \in S, \\ & \sigma(S) - a(S) \ge \varepsilon \quad \forall S \in \mathcal{S}_1^c \setminus \big\{\{n\}\big\} \text{ with } n \notin S, \\ & a(U) = v^T\sigma, \\ & a_n = 0. \end{aligned} \tag{15}$$

Now it easy to see that (15) is equivalent to (10), where $a_n = 0$ has been eliminated (without consequence) from the problem. This establishes Lemma 6.1.

References

Benjaafar S, Cooper WL, Kim J-S (2005) On the benefits of pooling in production-inventory systems. Manag Sci 51(4):548–5651

Chase R, Aquilano N (1992) Production and operations management, 6th edn. Irwin, Homewood

Cherikh M (2000) On the effect of centralization on expected profits in multi-location Newsboy problem. J Oper Res Soc 51:755–761

Dror M, Hartman BC, Chang W (2008) Core, concavity, and cost allocation for consolidation of inventory ordering. Working Paper, MIS, Eller College of Managament, University of Arizona (submitted for publication)

Eppen G (1979) Effects of centralization on expected costs in a multi-location newsboy problem. Manag Sci 25(5):498–501

Gerchak Y, Gupta D (1991) On apportioning costs to customers in centralized continuous review inventory systems. J Oper Manag 10:546–551

Gerchak Y, He Q-M (2003) On the relation between the benefits of risk pooling and the variability of demand. IIE Trans 35:1027–1031

Goemans M, Williamson D (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J ACM 42:1115–1145

Hartman BC (1994) Cooperative games and inventory cost allocation. PhD Thesis, Eller College of Management, University of Arizona

Hartman BC, Dror M (1996) Cost allocation in continuous review inventory models. Nav Res Logist 43:549–561

Hartman BC, Dror M (2003) Optimizing centralized inventory operations in a cooperative game theory setting. IIE Trans Oper Eng 35:243–257

Hartman BC, Dror M, Shaked M (2000) Cores of inventory centralization games. Games Econ Behav 31(1):26–49

Lee HL, Padmanabhan V, Whang S (1997) The bullwhip effect in supply chains. Sloan Manag Rev 38(3):93–102

Pataki G (1998) On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. Math Oper Res 23(2):339–358

Peleg B, Sudhölter P (2003) Introduction to the theory of cooperative games. Kluwer Academic, Dordrecht

Ruusunen J, Ehtamo H, Hamalainen R (1991) Dynamic cooperative electricity exchange in a power pool. IEEE Trans Syst Man Cybern 21(4):758–766

Schmeidler D (1969) The nucleolus of a characteristic function game. SIAM J Appl Math 17:1163–1170

Shapiro A (1982) Weighted minimum trace factor analysis. Psychometrika 47:243–264

Shapley L (1953) A value for n-person games. In: Contributions to the theory of games, vol 2. Princeton University Press, Princeton, pp 307–317

Shen Z-JM, Coullard C, Daskin MS (2003) A joint location-inventory model. Transp Sci 37(1):40–45

Wolkowicz H, Saigal R, Vandenberghe L (eds) (2000) Handbook of semidefinite programming. Kluwer Academic, Dordrecht

Young HP (1985a) Cost allocation: methods, principles, applications, Chap 1. Elsevier/North-Holland, Amsterdam, pp 3–29

Young HP (1985b) Monotonic solutions of cooperative games. Int J Game Theory 14:65–72