

Optimality Conditions and Exact Algorithms for Risk-Averse Bilevel Stochastic Linear Problems

Bruno M. Pacheco

December 13, 2024

Introduction

Bilevel stochastic problems can be seen as a generalization of the two-stage problems we have seen in class. In both cases, there are two decisions to be made: one before and another after the realization of a random variable. The difference lies in that bilevel stochastic programming does *not* assume that both decisions are made by the same agent. In turn, this difference leads to a bilevel problem because the two stages do not share the same objective.

The properties of bilevel stochastic linear problems have been studied in the foundational works by Burtsccheidt et al. (2020) and Claus (2021, 2022). The authors consider the more general risk-averse scenario, for which the risk-neutral case becomes a particular instance. They have presented proofs of the existence of optima and even optimality conditions for (classes of) bilevel problems in which the random variable appears in the right-hand side of the lower level (Burtsccheidt et al., 2020), in the lower level cost function (Claus, 2021) in a quadratic manner, or in both (Claus, 2022). Although those are solid results, their interpretation and applicability is not easy to grasp, as they are proposed for abstract problem classes and assume intricate properties from the components of the mathematical programming models (e.g., constraint functions, solution space, objective function).

The overarching goal of this project is to deepen the understanding of the

theoretical results for bilevel stochastic linear problems. The proposed approach is to explore the implications of these results for two classic textbook examples: the newsvendor problem and the multiproduct assembly problem. By proposing a bilevel variant of those problems and studying their theoretical properties following Burtsccheidt et al. (2020), I expect to make those results tangible for risk-averse bilevel stochastic linear problems. Finally, I expect that those applications lead to a clear idea of which exact algorithms can be used to solve the proposed problems, reaching a practical conclusion.

The Newsvendor Problem

As presented in the preliminary report, the newsvendor problem can be formulated as

$$\begin{aligned} \min_x \quad & cx + Q(x, z) \\ \text{s.t.} \quad & 0 \leq x \leq u, \end{aligned} \tag{1}$$

in which

$$\begin{aligned} Q(x, z) = \min_{y, w} \quad & -qy - rw \\ \text{s.t.} \quad & y \leq z \\ & y + w \leq x \\ & y, w \geq 0. \end{aligned} \tag{2}$$

The decision variables x , y , and w represent, respectively, the amount of newspaper initially bought, the amount of newspaper sold, and the amount of newspaper returned w . The problem is parameterized by the acquisition cost c , the newspaper capacity u , the demand z , the selling price q , and the return price r .

The traditional two-stage formulation comes from assuming that the demand comes from a random variable $z = Z(\omega)$, where ω belongs to a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, it is assumed that the realization of the random variable happens after the first decision (w.r.t. x), and before the second decision (w.r.t. y and w). Then, given a risk measure $\mathcal{R} : \mathcal{X} \rightarrow \mathbb{R}$, where \mathcal{X} is a linear subspace of all \mathcal{F} -measurable random variables, the two-stage

problem becomes

$$\begin{aligned} \min_x \quad & \mathcal{R}[cx + Q(x, Z)] \\ \text{s.t.} \quad & 0 \leq x \leq u. \end{aligned} \tag{3}$$

Note that if we assume that \mathcal{R} is translation invariant (which is a common property to expect, e.g., variance, value-at-risk), then $\mathcal{R}[cx + Q(x, Z)] = cx + \mathcal{R}[Q(x, Z)]$, which is true, for example, in the risk-neutral case $\mathcal{R} = \mathbb{E}$.

A Bilevel Variant

In this work, I will assume a slight variation of the original newsvendor problem in which the lower-level decision is made by a different agent, with a different objective. This may represent, for example, a scenario in which the newspaper acquisition is made by a middle-man, which has different selling and return margins than the newspaper salesperson. Instead of (1), we have, then,

$$\begin{aligned} \min_x \quad & f(x, z) = cx + \min \{-q_u y - r_u w : (y, w) \in \Psi(x, z)\} \\ \text{s.t.} \quad & 0 \leq x \leq u, \end{aligned} \tag{4}$$

in which $\Psi(x, z)$ represents the set of solutions to the lower-level problem, that is,

$$\begin{aligned} \Psi(x, z) = \arg \min_{y, w} \quad & -q_l y - r_l w \\ \text{s.t.} \quad & y \leq z \\ & y + w \leq x \\ & y, w \geq 0. \end{aligned} \tag{5}$$

Note that the costs differ, that is, the selling and returning costs for the upper level are q_u and r_u , resp., while they are q_l and r_l for the lower level.

The first interesting property to analyse, even before introducing the random variable, is that of the function f .

Lemma 1 (Burtsccheidt and Claus (2020), Lemma 17.2.1). Function f is real-valued and Lipschitz continuous $\forall x, z \geq 0$.

Proof. It is easy to see that $\forall x, z \geq 0$, the lower-level problem is feasible, and, thus, a solution to the minimization problem in $f(x, z)$ exists, which renders $f(x, z)$ a real-valued function.

Now, to demonstrate Lipschitz continuity, take any $x, z, x', z' \geq 0$ such that $f(x, z) \geq f(x', z')$. Then, take $(y', w') \in \Psi(x', z')$, which means that $f(x', z') = cx' - q_u y' - r_u w'$, and, thus, for any $(y, w) \in \Psi(x, z)$. On top of that, by Klatte and Thiere (1995, Theorem 4.2), we have that every point $(y', w') \in \Psi(x', z')$ can be expressed as

$$(y', w') = (y, w) + \Lambda \|(x, z) - (x', z')\| e$$

for some $(y, w) \in \Psi(x, z)$, a vector $e \in \mathbb{R}^2$ with $\|e\| \leq 1$, and some constant $\Lambda > 0$. Thus, assuming that $c \geq 0$ (which is indeed expected, as it represents a cost),

$$\begin{aligned} |f(x, z) - f(x', z')| &= f(x, z) - cx' + q_u y' + r_u w' \\ &\leq cx - q_u y - r_u w - cx' + q_u y' + r_u w' \\ &\leq c|x - x'| + \|(q_u, r_u)\| \|(y, w) - (y', w')\| \\ &\leq c|x - x'| + \|(q_u, r_u)\| \Lambda \|(x, z) - (x', z')\| \|e\| \\ &\leq L_f \|(x, z) - (x', z')\|, \end{aligned}$$

where $L_f = c + \Lambda \|(q_u, r_u)\|$. □

The Bilevel Stochastic Newsvendor

As for the two-stage problem, given a random demand Z and a risk measure \mathcal{R} , our bilevel newsvendor is interested in solving

$$\begin{aligned} \min_x \quad & \mathcal{R}[F(x)] \\ \text{s.t.} \quad & 0 \leq x \leq u, \end{aligned} \tag{6}$$

where $F(x) = f(x, Z)$ is a random variable parameterized by the upper-level decision x .

Before we dive in the properties of the random variable $F(x)$, it is necessary to lay out some definitions about Z . Let μ_Z be the Borel probability measure induced by Z . This means that, in face of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the probability that μ_Z associates to a given set $\{z_1, z_2, \dots\}$ of demand values

is equal to the probability of the subset of Ω that contains the respective realization values. In other words, if $\{\omega_1, \omega_2, \dots\} \in \Omega$ is such that $z_i = Z(\omega_i)$, for each $i = 1, 2, \dots$, then $\mu_Z(\{z_1, z_2, \dots\}) = \mathbb{P}\{\omega_1, \omega_2, \dots\}$. I will write this relationship as $\mu_Z = \mathbb{P} \circ Z^{-1}$, following Burtscheidt and Claus (2020).

This allows me to reformulate the result that presents the Lipschitz continuity of $F(x)$ for the bilevel newsvendor problem.

Lemma 2 (Burtscheidt and Claus (2020), Lemma 17.2.4^a). If μ_Z has finite moments of order $p \in [1, \infty)$, then $\exists L > 0$ such that

$$\|F(x) - F(x')\|_p \leq L|x - x'|, \quad \forall x, x' \geq 0,$$

i.e., $F(x)$ is Lipschitz continuous with respect to the L^p -norm $\|F(x)\|_p = \mathbb{E}[|F(x)|^p]^{1/p}$.

^aExcept the case for probability measures with finite moments of order $p = \infty$.

Proof. First, note that, given Lemma 1, $\forall x \geq 0$

$$\begin{aligned} (\|F(x)\|_p)^p &= \int_{\mathbb{R}^+} |f(x, z)|^p \mu_Z(dz) \\ &= \int_{\mathbb{R}^+} |f(x, z) - f(0, 0) + f(0, 0)|^p \mu_Z(dz) \\ &\leq \int_{\mathbb{R}^+} |f(x, z) - f(0, 0)|^p + |f(0, 0)|^p \mu_Z(dz) \\ &= |f(0, 0)|^p + \int_{\mathbb{R}^+} |f(x, z) - f(0, 0)|^p \mu_Z(dz) \\ &\leq |f(0, 0)|^p + \int_{\mathbb{R}^+} L_f^p \|(x, z)\|^p \mu_Z(dz) \\ &\leq |f(0, 0)|^p + L_f^p \int_{\mathbb{R}^+} (\|x\|^p + \|z\|^p) \mu_Z(dz) \\ &= |f(0, 0)|^p + L_f^p \|x\|^p + L_f^p \int_{\mathbb{R}^+} \|z\|^p \mu_Z(dz) < \infty, \end{aligned}$$

as μ_Z has finite moments of order p . □

References

- Burtscheidt, J., & Claus, M. (2020). Bilevel Linear Optimization Under Uncertainty. In S. Dempe & A. Zemkoho (Eds.), *Bilevel Optimization: Advances and Next Challenges* (pp. 485–511). Springer International Publishing. https://doi.org/10.1007/978-3-030-52119-6_17
- Burtscheidt, J., Claus, M., & Dempe, S. (2020). Risk-Averse Models in Bilevel Stochastic Linear Programming. *SIAM Journal on Optimization*, 30(1), 377–406. <https://doi.org/10.1137/19M1242240>
- Claus, M. (2021). On continuity in risk-averse bilevel stochastic linear programming with random lower level objective function. *Operations Research Letters*, 49(3), 412–417. <https://doi.org/10.1016/j.orl.2021.04.007>
- Claus, M. (2022). Existence of solutions for a class of bilevel stochastic linear programs. *European Journal of Operational Research*, 299(2), 542–549. <https://doi.org/10.1016/j.ejor.2021.12.004>
- Klatte, D., & Thiere, G. (1995). Error bounds for solutions of linear equations and inequalities. *Zeitschrift für Operations Research*, 41(2), 191–214. <https://doi.org/10.1007/BF01432655>