Two-Stage Stochastic Linear Programming:  
The Conditional Scenario Approach

C. Beltran-Royo*

19/01/2016

Abstract

In this paper we consider the Two-stage Stochastic Linear Programming (TSLP) problem with continuous random parameters. A common way to approximate the TSLP problem, generally intractable, is to discretize the random parameters into scenarios. Another common approximation only considers the expectation of the parameters, that is, the expected scenario. In this paper we introduce the conditional scenario concept which represents a midpoint between the scenario and the expected scenario concepts. The message of this paper is twofold: a) The use of scenarios gives a good approximation to the TSLP problem. b) However, if the computational effort of using scenarios results too high, our suggestion is to use conditional scenarios instead because they require a moderate computational effort and favorably compare to the expected scenario in order to model the parameter uncertainty.

Keywords: Stochastic programming, LP, conditional expectation, scenario, conditional scenario.

1 Introduction

As already pointed out, in this paper we assume that Two-stage Stochastic Linear Programming (TSLP) corresponds to Linear Programming (LP) problems with continuous random parameters and with two decision stages. The relevance for managerial purposes, properties, solution methods and applications of this problem can be found in surveys such as [5, 27] and in books [5, 28], among others. To address stochastic optimization problems different approaches can be used: robust optimization [4], chance constraint optimization [24, 29], sampling based methods [15] and scenario based optimization [5, 18], among others. In this paper we will focus on the last approach. On the other hand, the basic TSLP problem corresponds to a linear optimization problem with continuous variables, two stages and risk neutral. In this respect, several enhancements have been proposed in literature: TSLP problems with mixed-integer variables [1, 2], multi-stage versions of the TSLP problem [22, 30], TSLP problems with risk aversion [10, 23], etc. In this paper we deal with the basic TSLP problem, which for short, we call ‘the TSLP problem’.

The TSLP problem formulated in terms of a continuous random vector which accounts for all the uncertain parameters of the problem is, in general, numerically intractable. To address this difficulty one can approximate the original random vector by a discrete random vector with a finite number of realizations (the scenario tree). Thus, in the first step of scenario based optimization one calculates a representative and tractable scenario tree. See [9, 14, 16], among others. Once a representative scenario tree has been calculated, one formulates the so-called deterministic equivalent problem, usually, a

*cesar.beltran@urjc.es, Computer Science and Statistics, Rey Juan Carlos University, Madrid, Spain.
structured large scale LP problem. We will call it the two-stage linear recourse problem [5], or for short, the Recourse Problem (RP) to distinguish it from the TSLP problem, knowing that in literature, some times, both problems are used as synonyms. Otherwise said, in this paper, the RP problem is considered a discrete approximation of the TSLP problem. Notice that in the expression ‘RP problem’ we repeat the word ‘problem’. However we favor this expression instead of ‘R problem’ since in literature the acronym ‘RP’ is commonly used [5]. In the second step one solves the RP problem. General purpose optimization software, as for example CPLEX [17], can be used to solve RP instances of moderate size. However, in many cases one needs to use specialized approaches for stochastic programming as, for example, specialized interior point methods [6], decomposition methods such as Benders decomposition or Lagrangian relaxation [11, 26], parallel and/or grid computing [21], etc.

In general, the RP problem is considered a good approximation to the TSLP problem and to solve RP instances associated to real applications often requires a high computational effort. As an alternative to the RP problem one can use the Expected Value (EV) problem [5], where the random parameters of the TSLP problem are approximated by their expectations. That is, the EV problem approximates the TSLP problem by ignoring the parameter uncertainty. It is normally recognized that the EV problem requires a low computational effort but it is not an adequate approximation to the TSLP problem. Therefore, it seems that scenario based optimization (the RP problem) and deterministic optimization (the EV problem) represent two extreme choices regarding computational effort and ability to deal with uncertainty. A natural question can be raised: Would it be possible to propose a midpoint between these two approaches? That is, would it be possible an approach with a moderate computational effort and with a reasonable ability to deal with uncertainty? In this paper we give a positive unswear to this question by introducing the Conditional Expectation (CE) problem. As we will see, in the CE problem the random parameters of the TSLP problem are approximated by their conditional expectations. Thus, the CE problem improves the ability of the EV problem to deal with uncertainty by considering conditional expectations instead of expectations. On the other hand, the CE problem reduces the computational burden of the RP problem by considering conditional scenarios instead of scenarios (the conditional scenario concept will be introduced in Section 4). Of course, there is no such thing as a free lunch and the optimal CE solution is, in general, suboptimal for the RP problem but hopefully better than the optimal EV solution.

In summary, the objectives of this paper are: to introduce the CE problem, to study some of its theoretical properties and to show by example how one can effectively obtain good solutions for two-stage LP problems with uncertain parameters. With these objectives in mind, in Section 2 we describe the well known TSLP problem and the corresponding RP and EV approximations. In Section 3 we review the conditional expectation concept. In Section 4 we introduce the conditional scenario concept, to be used to formulate the CE problem in Section 5. In Section 6 some useful bounds related with problems RP, CE and EV are analyzed. In Section 7, a numerical example is used to illustrate the use of the CE problem. Section 8 concludes and outlines future research. Finally, in the Appendix we list and prove the theoretical results of Section 6.

2 The two-stage stochastic LP problem

2.1 Notation

Let us define the notation to be used in this paper. Vectors are assumed to be columns with transposes to indicate row vectors (e.g. $c$ is a column vector and $c^\top$ is a row vector). Random vectors are written in boldfaced font and the corresponding realizations are written in normal font. Examples: $\mathbf{b}$, $b$ and $\mathbf{\xi}, \xi$. Given a continuous random vector, say $\xi$, it will be approximated by the finite support random vectors $\tilde{\xi}$ and $\hat{\xi}$ which are based on scenarios and conditional scenarios, respectively. The tilde symbol $\tilde{}$ is used for scenarios, say $\tilde{\xi}$, as well as for the corresponding random vector $\tilde{\xi}$ and probability value $\tilde{p} = P(\tilde{\xi} = \xi)$. The hat symbol $\hat{}$ is used for conditional scenarios, say $\hat{\xi}$, as well as for the corresponding random vector $\hat{\xi}$ and probability value $\hat{p} = P(\hat{\xi} = \xi)$. As usual, subscripts are used for components.
of vectors, matrices, etc. Examples:

\[
a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (a_j)_{j \in J}
\]

\[
A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (a_{ij})_{i \in I, j \in J}
\]

for given index sets \( I \) and \( J \).

Indexes:

- \( e \) Realizations of random variables \( e \in E = \{1, \ldots, E\} \)
- \( r \) Components of random vectors \( r \in R = \{1, \ldots, R\} \)
- \( s \) Scenarios \( s \in S = \{1, \ldots, S\} \)
- \( re \) Index pair for conditional scenarios \( re \in R \times E = \{1, \ldots, E_r\} \)

Parameters:

- \( \xi \) Continuous random vector which accounts for all the random parameters of an LP problem
- \( \bar{\xi} \) Expectation of \( \xi \), that is, \( \bar{\xi} = E[\xi] \)
- \( \tilde{\xi} \) Finite support random vector which approximates \( \xi \)
- \( \hat{\xi} \) Scenario or realization of the random vector \( \tilde{\xi} \)
- \( \hat{p} \) Component \( r \) of the random vector \( \hat{\xi} \)
- \( \hat{\xi}_{re} \) Realization of the random variable \( \hat{\xi}_{r} \)
- \( \hat{\xi}^r \) Random vector which approximates \( \tilde{\xi} \) by conditional expectation, that is, \( \hat{\xi}^r = E[\tilde{\xi} | \tilde{\xi}_{r}] \)
- \( \hat{\xi}^{re} \) Conditional scenario or realization of the random vector \( \hat{\xi}^r \)
- \( \hat{p}^{re} \) Probability of \( \hat{\xi}^{re} \), that is, \( \hat{p}^{re} = P(\hat{\xi}^r = \hat{\xi}^{re}) \)

Other symbols:

- \( c^T x \) Scalar product of vectors \( c \) and \( x \), that is, \( c^T x = \sum_{j \in J} c_j x_j \)
- \( \text{vec}(u_1, \ldots, u_J) \) Vector of vectors:
  \[
  \text{vec}(u_1, \ldots, u_J) = \begin{pmatrix} u_1 \\ \vdots \\ u_J \end{pmatrix}, \text{ for any set of vectors } \{u_1, \ldots, u_J\}
  \]
- \( \text{vec}(A) \) Matrix \( A \) in vector format:
  \[
  \text{vec}(A) = \begin{pmatrix} a_{s1} \\ \vdots \\ a_{sJ} \end{pmatrix} \text{ where } A = (a_{s1} \cdots a_{sJ})
  \]
  and \( a_{sj} \) is the \( j \)th column of matrix \( A \) for all \( j \in J \)
- \( \xi \sim N_R(\mu, \Sigma) \) \( \xi \) is a multinormal random vector with dimension \( R \), mean vector \( \mu \) and covariance matrix \( \Sigma \)
  (for the case \( R = 1 \) we simply write \( \xi \sim N(\mu, \sigma) \))
2.2 Problem formulation

As already pointed out, in this paper we consider the Two-stage Stochastic LP problem (TSLP) where some or all the entries of \(c_2, B_2, A_2\) and \(b_2\) are continuous random variables. This stochastic problem can be written as follows

\[
\begin{align*}
\min_x & \quad z_{TSLP} = c_1^T x_1 + \mathbb{E}[c_2^T x_2] \\
\text{s.t.} & \quad A_1 x_1 = b_1 \\
& \quad B_2 x_1 + A_2 x_2 = b_2 \quad \text{w.p.} 1 \\
& \quad x_2 \geq 0 \quad \text{w.p.} 1 \\
& \quad x_1 \geq 0.
\end{align*}
\]

where ‘w.p. 1’ stands for ‘with probability 1’. For example, in (3) it means that \(\xi\) one single vector. Since this problem is, in general, intractable one solves an approximation based on a deterministic equivalent version:

\[
\begin{align*}
\min_{\bar{x}} & \quad z_{RP} = c_1^T \bar{x}_1 + \mathbb{E}[\bar{c}_2^T \bar{x}_2(\bar{\xi})] \\
\text{s.t.} & \quad A_1 \bar{x}_1 = b_1 \\
& \quad B_2 \bar{x}_1 + A_2 \bar{x}_2(\bar{\xi}) = \bar{b}_2 \quad \text{w.p.} 1 \\
& \quad \bar{x}_1 \geq 0, \quad \bar{x}_2(\bar{\xi}) \geq 0 \quad \text{w.p.} 1
\end{align*}
\]

In this paper we call the discrete approximation to the TSLP problem based on \(\bar{\xi}\), the two-stage linear Recourse Problem (RP). Thus, the RP problem can be written as follows (stochastic version):

\[
\begin{align*}
\min_{\bar{x}} & \quad z_{RP} = c_1^T \bar{x}_1 + \mathbb{E}[\bar{c}_2^T \bar{x}_2(\bar{\xi})] \\
\text{s.t.} & \quad A_1 \bar{x}_1 = b_1 \\
& \quad B_2 \bar{x}_1 + A_2 \bar{x}_2(\bar{\xi}) = \bar{b}_2 \quad \text{w.p.} 1 \\
& \quad \bar{x}_1 \geq 0, \quad \bar{x}_2(\bar{\xi}) \geq 0 \quad \text{w.p.} 1
\end{align*}
\]

where \(\bar{\xi} = \text{vec}(\bar{c}_2, B_2, A_2, \bar{b}_2)\) accounts for a finite support approximation to the continuous random vector \(\xi\) (the random parameters have been written in boldface to distinguish them from the deterministic ones, namely, \(c_1, A_1\) and \(b_1\)). For numerical purposes the RP problem is written in the so-called deterministic equivalent version:

\[
\begin{align*}
\min_{\tilde{x}} & \quad z_{RP} = c_1^T \tilde{x}_1 + \sum_{s \in S} \tilde{p}^s \tilde{c}_2^T \tilde{x}_2^s \\
\text{s.t.} & \quad A_1 \tilde{x}_1 = b_1 \\
& \quad B_2^s \tilde{x}_1 + A_2^s \tilde{x}_2^s = \tilde{b}_2^s \\
& \quad \tilde{x}_1 \geq 0, \quad \tilde{x}_2^s \geq 0 \quad s \in S
\end{align*}
\]

such that \(\tilde{\xi} = \text{vec}(\tilde{c}_2, B_2^s, A_2^s, \tilde{b}_2^s)\) for all \(s \in S\).

In cases where the RP problem results computationally unaffordable, one can approximate it by the so-called Expected Value (EV) problem. More precisely, the EV problem corresponds to approximate \(\xi\) by \(\bar{\xi} = \mathbb{E}[\xi] = \text{vec}(\bar{c}_2, B_2, A_2, b_2)\) and can be stated as

\[
\begin{align*}
\min_{\bar{x}} & \quad z_{EV} = c_1^T \bar{x}_1 + \bar{c}_2^T \bar{x}_2 \\
\text{s.t.} & \quad A_1 \bar{x}_1 = b_1 \\
& \quad B_2 \bar{x}_1 + A_2 \bar{x}_2 = \bar{b}_2 \\
& \quad \bar{x}_1 \geq 0, \quad \bar{x}_2 \geq 0.
\end{align*}
\]

Notice that notation \(\bar{x}\) and \(\tilde{x}\) is used to distinguish the decision vectors of the EV problem from the RP ones.
3 Conditional expectation

Given that the conditional expectation problem, to be introduced in Section 5, is based on the conditional expectation, let us first review this concept. As it is well known, given a random vector \( \xi = (\xi_1 \ldots \xi_R)^T \), from an intuitive point of view, its expectation \( E[\xi] \) corresponds to the long-run average vector of repetitions of the experiment \( \xi \) represents. In the case of having dependent components in \( \xi \), it can be advantageous to use the possible knowledge of one component to compute the expectation of the whole random vector. For example, if one observes that \( \xi_1 = 100 \), the conditional expectation of \( \xi \) given \( \xi_1 = 100 \), that is, \( E[\xi | \xi_1 = 100] \), will be a better picture of the long-run average vector of repetitions of the experiment \( \xi \) represents, than \( E[\xi] \). In this case one writes \( E[\xi | \xi_1 = 100] \) or for short, \( E[\xi | \xi_r] \) for any \( r \in \mathcal{R} = \{1, \ldots, R\} \). Instead of only considering an observation \( \xi_r \), one can consider the random variable itself \( \xi_r \) and compute \( E[\xi | \xi_r] \), the conditional expectation of \( \xi \) given \( \xi_r \). For short, in this paper we denote it by \( \xi_r \) and call it the \( r \)th conditional expectation of \( \xi \), for any \( r \in \mathcal{R} \).

We restrict ourselves to the multinormal distribution to simplify the exposition, knowing that the conditional expectation is a well known concept in Statistics, which can be computed for any probability distribution. The multinormal distribution [13] is characterized by a random vector \( \xi \) with mean \( \mu \) and covariance matrix \( \Sigma > 0 \) (positive definite) and by the corresponding probability density function

\[
f(\xi) = |2\pi \Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\xi - \mu)^\top \Sigma^{-1} (\xi - \mu) \right\}, \quad \xi \in \mathbb{R}^R,
\]

where \( | \cdot | \) stands for the determinant. We write \( \xi \sim N_R(\mu, \Sigma) \).

**Proposition 1.** (Definition 5.1) Let us consider a multinormal random vector \( \xi = (\xi_1 \ldots \xi_R)^T \sim N_R(\mu, \Sigma) \). Then the conditional expectation of \( \xi \) given \( \xi_r \) can be calculated as follows:

\[
E[\xi | \xi_r] = \mu + \frac{\xi_r - \mu_r}{\sigma_r^2} \Sigma_{sr}\]

\[r \in \mathcal{R}\]  \quad (14)

where \( \Sigma_{sr} \) is the \( r \)th column of the covariance matrix \( \Sigma \).

Notice that so far we have only given an intuitive interpretation of the conditional expectation and a formula to compute it in the case of a multinormal random vector. The definition of the conditional expectation and the formula to compute it for a general distribution can be found in [8, 13].

**Example 1.** (Conditional expectation of a multinormal random vector)

Let us consider the multinormal random vector \( \xi = (\xi_1 \xi_2)^T \sim N_2(\mu, \Sigma) \) such that

\[
\mu = \begin{pmatrix} 100 \\ 200 \end{pmatrix}^T, \quad \Sigma = \begin{pmatrix} 400 & 480 \\ 480 & 1600 \end{pmatrix}.
\]

Our objective is to compute its 1st conditional expectation \( \xi_1 \). According to Proposition 1, we have that

\[
\xi_1^1 = E[\xi | \xi_1] = \mu + \frac{\xi_1 - \mu_1}{\sigma_1^2} \Sigma_{s1} = \begin{pmatrix} 100 \\ 200 \end{pmatrix} + \frac{\xi_1 - 100}{400} \begin{pmatrix} 400 \\ 480 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ 80 + 1.2\xi_1 \end{pmatrix},
\]

where \( \xi_1 \sim N(\mu_1, \sigma_1) \) with \( \mu_1 = 100 \) and \( \sigma_1 = 20 \). Notice that in this case \( \xi_1 \) is a straight line (see Figure 1). To interpret this result let us assume, for example, that we have observed \( \xi_1 = 50 \).
Then, the expected value of $\xi_2$, given $\xi_1 = 50$, is 140. This value is different from the expected value of $\xi_2$, which is 200 units. Analogously, we could compute the 2nd conditional expectation $\xi_2 = (40 + 0.3\xi_2, \xi_2)^T$ where $\xi_2 \sim N(\mu_2, \sigma_2)$ with $\mu_2 = 200$ and $\sigma_2 = 40$ (with our notation, $\sigma_2$ is the standard deviation, not to be confused with the corresponding variance $\sigma_2^2$).

4 Conditional scenarios

In this section we introduce the conditional scenario concept. Given a continuous random vector $\xi = (\xi_1 \ldots \xi_R)^T$, we can approximate it by a discrete random vector $\hat{\xi} = (\hat{\xi}_1 \ldots \hat{\xi}_R)^T$, with finite support $S_{\hat{\xi}} = \{\hat{x}_s\}_{s \in S}$ and the corresponding probabilities $\{\hat{p}_s\}_{s \in S}$ with $S = \{1, \ldots, S\}$. In this context, each realization $\hat{x}_s$ is called a scenario. One basic way to discretize $\xi$ is to discretize each component $\xi_r$ into, say $E$ points, and then consider the Cartesian product of these component discretizations. The result is a grid of $E^R$ points in $\mathbb{R}^R$. Otherwise said, the Cartesian discretization of $\xi$ has $E^R$ scenarios. This exponential number of scenarios produces a Recourse Problem (RP) which is huge, thus difficult to solve. In this context some technique to consider a reduced number of representative scenarios, as for example, the moment matching method [16, 19] or the Sample Average Approximation (SAA) method [15, 20], among others, is usually used.

As an alternative, one may approximate $\xi$ considering the $r$th conditional expectation $\hat{\xi}^r = \mathbb{E}[\hat{\xi} | \xi_e]$ for any $r \in R$, and then formulate the Conditional Expectation (CE) problem in terms of $\hat{\xi}^r$ (this problem will be defined in next section). The conditional expectation $\hat{\xi}^r = (\hat{\xi}_1^r \ldots \hat{\xi}_R^r)^T$, is a random vector with finite support $S_{\hat{\xi}^r} = \{\hat{\xi}^r_e\}_{e \in E}$ and the corresponding probability values $\hat{p}^r_e = P(\hat{\xi}^r = \hat{\xi}^r_e)$ for all $e \in E_r = \{1, \ldots, E_r\}$. Each realization $\hat{\xi}^r_e$ is called a conditional expectation scenario or, for short, conditional scenario. Notice that we use symbol $\hat{x}_s$ with $s \in S$ for scenarios and symbol $\hat{\xi}^r_e$ with $e \in E_r$ for conditional scenarios.

Now let us put things in perspective for a more clear exposition. The introduction of the conditional scenario concept has been as follows: the TSLP problem (1)–(5) has been stated in terms of the continuous random vector $\xi$. Since this problem is in general intractable, one can approximate $\xi$ by $\hat{\xi}$ and define the RP problem which is based on scenarios (realizations of $\xi$). Given that this problem, in some cases may require a high computational effort, one can approximate $\xi$ by the set of conditional expectation vectors $\{\xi^r\}_{r \in R}$ and define the CE problem which is based on conditional scenarios (realizations of $\xi^1, \ldots, \xi^R$), as we will see in next section. In summary, a ‘discretization + conditional expectation’ scheme has been applied to obtain the chain $\xi \rightarrow \hat{\xi} \rightarrow \{\hat{\xi}^r\}_{r \in R}$. That is, we have derived the conditional scenario concept from the scenario concept.

In fact, one can define the scenario and conditional scenario concepts independently. To do so, it would be enough to apply a ‘conditional expectation + discretization’ scheme to obtain the chain $\xi \rightarrow \{\xi^r\}_{r \in R} \rightarrow \{\hat{\xi}^r\}_{r \in R}$. That is one approximates the continuous random vector $\xi$ by the set of (continuous) conditional expectation vectors $\{\xi^r\}_{r \in R}$ and then discretize them into $\{\hat{\xi}^r\}_{r \in R}$ in order to define the CE problem. The two schemes just described to obtain the conditional scenarios are equivalent in the sense that they would produce the same set of conditional scenarios for sufficiently fine discretizations in both schemes (with a large enough number of discretization points).

In this paper we will concentrate on the scheme ‘discretization + conditional expectation’ in order to relate and compare problems RP, CE and EV. However, in this and in the previous sections it is more handy to use the alternative scheme ‘conditional expectation + discretization’ to illustrate the conditional scenario concept in the case of the multinormal distribution. Although one can define conditional scenarios for any probability distribution, the advantage of the multinormal distribution is that formula (14) allows for a straightforward calculation of the conditional scenarios. That is, the use of the multinormal distribution simplifies our exposition.

**Method 1.** *(Conditional scenarios of a multinormal random vector)*

- **Scheme:  conditional expectation + discretization.*
• **Objective**: To approximate a multinormal random vector \( \xi \) by a finite set of conditional scenarios.

• **Input**: \( \xi = (\xi_1 \ldots \xi_R)^T \sim N_R(\mu, \Sigma) \), a multinormal random vector. \( E_r \), the number of conditional scenarios associated to coordinate \( r \), for all \( r \in R \).

• **Output**: The set of random vectors \( \{\hat{\xi}^r\}_{r \in R} \), where for each \( r \in R \), we have the support \( S_{\hat{\xi}^r} = \{\hat{\xi}^{re}\}_{e \in E_r} \) and the corresponding probabilities \( \{\hat{p}^{re}\}_{e \in E_r} \), such that \( \hat{p}^{re} = P(\xi^r = \hat{\xi}^{re}) \) for all \( e \in E_r = \{1, \ldots, E_r\} \). Each \( \hat{\xi}^{re} \) is called a conditional scenario.

• **Steps**: For all \( r \in R \),

1) Compute the \( r \)th conditional expectation (Proposition 1):

\[
\hat{\xi}^r = \mathbb{E}[\xi | \xi_r] = \mu + \frac{\xi_r - \mu_r}{\sigma_r^2} \Sigma_{sr}.
\]

2) Discretize the continuous random variable \( \xi_r \) into the random variable \( \hat{\xi}_r \) which has finite support \( \{\hat{\xi}_{re}\}_{e \in E_r} \) and the corresponding probabilities \( \{\hat{p}_{re}\}_{e \in E_r} \).

3) By using the discrete random variable \( \hat{\xi}_r \), discretize the continuous random vector \( \xi^r \) into the random vector \( \hat{\xi}^{re} \) which has finite support \( \{\hat{\xi}^{re}\}_{e \in E_r} \) and the corresponding probabilities \( \{\hat{p}^{re}\}_{e \in E_r} \). More precisely:

\[
\hat{\xi}^{re} = \mathbb{E}[\xi | \hat{\xi}_{re}] = \mu + \frac{\hat{\xi}_{re} - \mu_r}{\sigma_r^2} \Sigma_{sr}
\]

\[
\hat{p}^{re} = \hat{p}_{re}
\]

for all \( e \in E_r \).

**Example 2.** (Conditional scenarios of a multinormal random vector) Let us consider again \( \xi \), the multinormal random vector of the Example 1. In order to illustrate the previous method, let us approximate \( \xi \) by six conditional scenarios derived from the 1st conditional expectation. In the first step of the method one computes the 1st conditional expectation (already done in Example 1):

\[
\hat{\xi}^1 = \mathbb{E}[\xi | \xi_1] = \begin{pmatrix} \xi_1 \\ 80 + 1.2\xi_1 \end{pmatrix},
\]

(15)

where \( \xi_1 \sim N(\mu_1, \sigma_1) \) with \( \mu_1 = 100 \) and \( \sigma_1 = 20 \). In the second step one discretizes the random variable \( \xi_1 \) into six representative points, as for example,

\[
\hat{\xi}_{1,1} = 50 \quad \hat{p}_{1,1} = 0.0176 \\
\hat{\xi}_{1,2} = 70 \quad \hat{p}_{1,2} = 0.1297 \\
\hat{\xi}_{1,3} = 90 \quad \hat{p}_{1,3} = 0.3527 \\
\hat{\xi}_{1,4} = 110 \quad \hat{p}_{1,4} = 0.3527 \\
\hat{\xi}_{1,5} = 130 \quad \hat{p}_{1,5} = 0.1297 \\
\hat{\xi}_{1,6} = 150 \quad \hat{p}_{1,6} = 0.0176.
\]

This discretization has been obtained by dividing the interval \( \mathcal{I} = [\mu_1 - 3\sigma_1, \mu_1 + 3\sigma_1] \) into six subintervals of equal length and by taking the midpoint of each subinterval as the corresponding discretization point \( \hat{\xi}_{1e} \) for \( e = 1, \ldots, 6 \). The probability values have been computed as:

\[
\hat{p}_{1e} = \frac{f(\hat{\xi}_{1e})}{\sum_{e=1}^{6} f(\hat{\xi}_{1e})}, \quad e = 1, \ldots, 6
\]

where \( f \) is the probability density function of random variable \( N(\mu_1, \sigma_1) \). This is a simple discretization way and, of course, other approaches could be used. In the third step one computes the conditional
Figure 1: Equiprobability contour ellipses of the multinormal random vector $\xi$ of Example 2. The dashed and the solid lines correspond to the first ($\xi^1$) and the second ($\xi^2$) conditional expectations, respectively. The dots represent the twelve conditional scenarios and the square represents the ‘expected scenario’ $\mu$.

scenarios combining the previous discretization with (15). For example the first conditional scenario $\hat{\xi}_{11}$ can be computed as follows:

$$\hat{\xi}^{1,1} = \mathbb{E}[\xi | \tilde{\xi}_{1,1}] = \left( \begin{array}{c} \tilde{\xi}_{1,1} \\ 80 + 1.2\tilde{\xi}_{1,1} \end{array} \right) = \left( \begin{array}{c} 50 \\ 80 + 1.2 \cdot 50 \end{array} \right) = \left( \begin{array}{c} 50 \\ 140 \end{array} \right)$$

$$\hat{\tilde{p}}^{1,1} = \tilde{p}_{1,1} = 0.0176$$

In summary we obtain:

$$\hat{\xi}^{1,1} = \left( \begin{array}{c} 50 \\ 140 \end{array} \right)^T \quad \hat{\tilde{p}}^{1,1} = \tilde{p}_{1,1} = 0.0176$$

$$\hat{\xi}^{1,2} = \left( \begin{array}{c} 70 \\ 164 \end{array} \right)^T \quad \hat{p}^{1,2} = \tilde{p}_{1,2} = 0.1297$$

$$\hat{\xi}^{1,3} = \left( \begin{array}{c} 90 \\ 188 \end{array} \right)^T \quad \hat{p}^{1,3} = \tilde{p}_{1,3} = 0.3527$$

$$\hat{\xi}^{1,4} = \left( \begin{array}{c} 110 \\ 212 \end{array} \right)^T \quad \hat{p}^{1,4} = \tilde{p}_{1,4} = 0.3527$$

$$\hat{\xi}^{1,5} = \left( \begin{array}{c} 130 \\ 236 \end{array} \right)^T \quad \hat{p}^{1,5} = \tilde{p}_{1,5} = 0.1297$$

$$\hat{\xi}^{1,6} = \left( \begin{array}{c} 150 \\ 260 \end{array} \right)^T \quad \hat{p}^{1,6} = \tilde{p}_{1,6} = 0.0176$$

The approximation of $\xi$ by six conditional scenarios derived from the 2nd conditional expectation $\hat{\xi}^2$ could be done analogously. In total we would obtain $6 \cdot 2$ conditional scenarios versus the $6^2$ Cartesian scenarios $\{\xi^e\}_{e=1}^{36}$ that would be obtained as the Cartesian product $\{\xi_{1e}\}_{e=1}^6 \times \{\xi_{2e}\}_{e=1}^6$. The 12 conditional scenarios are represented in Figure 1 as equidistant points along the conditional expectations $\xi_1$ and $\xi_2$. As we will see in next section, the conditional expectation problem is formulated in terms of conditional scenarios, in contrast with the EV problem, which is formulated in terms of the ‘expected scenario’ $\mu$ (the square in Figure 1).

Considering that by definition the random vector $\xi^r = \mathbb{E}[\xi | \xi_r]$ is a transformation of the random variable $\xi_r$, the number of the corresponding conditional scenarios is equal to the number of points used to discretize $\xi_r$ into $\tilde{\xi}_r$, say $E$. Otherwise said, we have a conditional scenario $\hat{\xi}^{re}$ per each discretization point $\tilde{\xi}_{re}$. Taking into account that the vectors to be discretized are $\xi^1, \ldots, \xi^R$, the total number of conditional scenarios is $E \cdot R$. Therefore, problems RP and CE require $E^R$ (Cartesian) scenarios and $E \cdot R$ conditional scenarios, respectively. Of course, the reduced size of the CE problem
also reduces the performance of the optimal CE solution which, in general, is suboptimal for the RP problem but hopefully better than the optimal EV solution. Regarding the potentially large size of the RP problem, in real applications one normally uses some scenario reduction technique, as for example the moment matching method or the SAA method, to obtain tractable RP instances, as already pointed out in this section.

5 The conditional expectation problem

Next we define the Conditional Expectation (CE) problem as the approximation to the Recourse Problem (RP) based on the approximation of the random vector \( \tilde{\xi} \) by a set of \( R \) random vectors, that is, \( \{ \tilde{\xi}^r \}_{r \in \mathcal{R}} \), each one taken with probability \( 1/R \). Otherwise said, to define the CE problem one approximates the random vector \( \hat{\xi} \) of dimension \( R \) by \( \hat{\xi}^r \) a random vector of dimension \( R + 1 \). Furthermore, the set of random vectors \( \{ \hat{\xi}^r \}_{r \in \mathcal{R}} \) is a family of optimal approximations to \( \hat{\xi} \) as seen in Proposition 2. Therefore, the CE problem has the appealing property of being an optimal approximation to the RP problem (in the sense of Proposition 2).

In summary, the CE problem is based on the approximation of \( \hat{\xi} \) by \( \{ \hat{\xi}^r \}_{r \in \mathcal{R}} \) and can be stated as
follows (stochastic version):

$$\min_{\hat{x}} z_{CE} = c_1^T \hat{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[\hat{c}_2^T \hat{x}_2(\tilde{\xi}_r)]$$  \hspace{1cm} (16)

s.t. \hspace{0.5cm} A_1 \hat{x}_1 = b_1  \hspace{1cm} (17)

$$\hat{B}_1^e \hat{x}_1 + \hat{A}_1^e \hat{x}_2(\xi_1) = \hat{b}_1^e \hspace{1cm} w.p.1$$  \hspace{1cm} (18)

$$\hat{B}_2^e \hat{x}_1 + \hat{A}_2^e \hat{x}_2(\xi_2) = \hat{b}_2^e \hspace{1cm} w.p.1$$  \hspace{1cm} (19)

$$...$$  \hspace{1cm} (20)

$$\hat{B}_2^R \hat{x}_1 + \hat{A}_2^R \hat{x}_2(\tilde{\xi}_R) = \hat{b}_2^R \hspace{1cm} w.p.1$$  \hspace{1cm} (21)

$$\hat{x}_1 \geq 0, \hspace{0.5cm} \hat{x}_2(\tilde{\xi}_r) \geq 0 \hspace{1cm} w.p.1, \hspace{0.2cm} r \in \mathcal{R},$$  \hspace{1cm} (22)

where

$$\tilde{\xi} = \text{vec}(\tilde{c}_2, \tilde{B}_2, \tilde{A}_2, \tilde{b}_2)$$

$$\hat{c}_2^r = \mathbb{E}[\hat{c}_2 | \tilde{\xi}_r] \hspace{1cm} r \in \mathcal{R}$$

$$\hat{B}_2^r = \mathbb{E}[\hat{B}_2 | \tilde{\xi}_r] \hspace{1cm} r \in \mathcal{R}$$

$$\tilde{A}_2^r = \mathbb{E}[\tilde{A}_2 | \tilde{\xi}_r] \hspace{1cm} r \in \mathcal{R}$$

$$\hat{b}_2^r = \mathbb{E}[\hat{b}_2 | \tilde{\xi}_r] \hspace{1cm} r \in \mathcal{R}$$

$$\hat{\xi}^r = \mathbb{E}[\hat{\xi} | \tilde{\xi}_r] = \text{vec}(\hat{c}_2^r, \hat{B}_2^r, \tilde{A}_2^r, \hat{b}_2^r) \hspace{1cm} r \in \mathcal{R}.$$

Finally, we state the CE problem in the deterministic equivalent version:

$$\min_{\hat{x}} z_{CE} = c_1^T \hat{x}_1 + \frac{1}{R} \sum_{re \in R\mathcal{E}_r} \hat{p}^{re} \hat{c}_2^{r}\hat{x}_2^{re}$$  \hspace{1cm} (23)

s.t. \hspace{0.5cm} A_1 \hat{x}_1 = b_1  \hspace{1cm} (24)

$$\hat{B}_1^{re} \hat{x}_1 + \hat{A}_1^{re} \hat{x}_2^{re} = \hat{b}_1^{re} \hspace{1cm} e \in \mathcal{E}_1$$  \hspace{1cm} (25)

$$\hat{B}_2^{re} \hat{x}_1 + \hat{A}_2^{re} \hat{x}_2^{re} = \hat{b}_2^{re} \hspace{1cm} e \in \mathcal{E}_2$$  \hspace{1cm} (26)

$$...$$  \hspace{1cm} (27)

$$\hat{B}_2^{Re} \hat{x}_1 + \hat{A}_2^{Re} \hat{x}_2^{Re} = \hat{b}_2^{Re} \hspace{1cm} e \in \mathcal{E}_R$$  \hspace{1cm} (28)

$$\hat{x}_1 \geq 0, \hspace{0.5cm} \hat{x}_2^{re} \geq 0 \hspace{1cm} re \in \mathcal{R}\mathcal{E}_r.$$  \hspace{1cm} (29)

Notice that these two versions are equivalent. The deterministic equivalent version is used for numerical purposes (see Section 7) and the stochastic one for theoretical analysis (see Section 9).

### 6 Useful bounds

In this section we give some bounds which can be used to assess the quality of the optimal EV and CE solutions. Normally, problems EV and CE are solved when the computational effort to solve the RP problem results too high. In this case, one cannot assess the quality of the optimal EV and CE solutions by using the optimal RP cost, that is, one cannot compute the EV and CE optimality gaps. If, at least, one has a lower bound for the RP optimal cost, a worst-case optimality gap can be computed for the optimal EV and CE solutions (see Section 7.3 for details). In what follows the values $z_{EV}^*$, $z_{CE}^*$ and $z_{RP}^*$ stand for the optimal cost of problems EV, CE and RP, respectively. We give some theoretical results whose proofs are in the Appendix (Section 9).

**Proposition 3.** Let us consider the RP problem (6)–(9) where some or all the components of $B_2$ and/or $b_2$ are stochastic (the other parameters being deterministic). Then

$$z_{EV}^* \leq z_{CE}^* \leq z_{RP}^*.$$
Proposition 4. Let us consider the RP problem (6)–(9) where some or all the components of \( c_2 \) are stochastic (the other parameters being deterministic). Then

\[ z^*_\text{RP} \leq z^*_\text{CE} \leq z^*_\text{EV}. \]

To approximate the RP problem by the EV problem has three main drawbacks. First, as one would expect, the optimal EV decision \( \bar{x}^*_1 \) is suboptimal for the RP problem (in this section we will assume that \( \bar{x}^*_1 \) is feasible for the RP problem). Second, the optimal EV cost \( z^*_\text{EV} \) usually is misleading, since it does not correspond to the true expected cost associated to the optimal EV solutions \( \bar{x}^*_1 \). To know the true expected cost one can compute the so-called expected result of using the EV solution (EEV) [5]. It can be computed by solving the EEV problem which corresponds to the RP problem with the additional constraint \( \tilde{x}_1 = \bar{x}^*_1 \). Third, the EV problem does not take parameter uncertainty, if any, into account.

The approximation of the RP problem by the CE problem, although avoids the third drawback of the EV problem, it also has the first two drawbacks. For this reason, we define the expected result of using the CE solution (ECE) as the counterpart of the EEV, that is, the RP problem with the additional constraint \( \tilde{x}_1 = \hat{x}^*_1 \) (in this section we will also assume that \( \hat{x}^*_1 \) is feasible for the RP problem).

Proposition 5. Let us consider the RP problem (6)–(9) where some or all the components of \( c_2 \) are stochastic (the other parameters being deterministic). Then

\[ z^*_\text{RP} \leq z^*_\text{EEV} \leq z^*_\text{EV} \quad \text{and} \quad z^*_\text{RP} \leq z^*_\text{ECE} \leq z^*_\text{CE}. \]

The previous bounds can only be applied when one restricts the set of stochastic parameters to \( B_2, b_2 \) or to \( c_2 \). Next result applies for all the cases.

Proposition 6. Let us consider the RP problem (6)–(9) where some or all the second stage parameters are stochastic (\( c_1, A_1 \) and \( b_1 \) are always supposed to be deterministic). Then

\[ z^*_\text{RP} \leq z^*_\text{EEV} \quad \text{and} \quad z^*_\text{RP} \leq z^*_\text{ECE}. \]

Remark 1. A particularly useful case is the RP problem where some or all the components of \( B_2 \) and/or \( b_2 \) are stochastic (the other parameters being deterministic). In this case we have that

\[ z^*_\text{CE} \leq z^*_\text{RP} \leq z^*_\text{ECE} \quad \text{and} \quad z^*_\text{CE} \leq z^*_\text{RP} \leq z^*_\text{EEV}, \]

which allows to compute the worst-case optimality gaps for the optimal EV and CE solutions (notice that we use \( z^*_\text{CE} \) in both cases, since by Proposition 3 it is a tighter bound than \( z^*_\text{EV} \) for the optimal RP cost \( z^*_\text{RP} \). In next section we will illustrate this case.

7 Numerical example

In this section we use a feed manufacturer problem which can be modelled as a Two-stage Stochastic LP problem (TSLP) with uncertain left-hand side. This problem has been inspired by the farmer’s problem in [5] (Section 1.1.a). As already pointed out in the Introduction, the Recourse Problem (RP) is, in general, a good choice to approximate the TSLP problem. However, in some cases the computational effort to solve the RP problem may result too high, which is our assumption here. That is, let us assume that the practitioner considers that the computational effort to solve the RP problem results unaffordable and therefore he needs some approach which requires less computational effort. With this assumption in mind, in this section we will compare problems EV and CE. We also assume that the practitioner is aware that better solutions could be found by solving the corresponding RP problem. Computations have been conducted on a PC under Windows 7 (64 bits), with a processor Intel Core i5, 2.67GHz and 8 GB of RAM. The LP problems have been solved by CPLEX 12.6 with default parameters.
7.1 The EV feed manufacturer problem

**Example 3.** *(The EV feed manufacturer problem: expected yields)*

Consider a feed manufacturer who specializes in raising and manufacturing several types of feed ingredients as for example grain mixes, orange rinds, beet pulps, etc. The grain mixes include corn, soybeans, sorghum, oats, and barley, among others. In total he plants fifty types of crops on his 1,700 acres of land. Based on experience the manufacturer knows, for each crop \( j \): a) That the mean yield on his land is \( \bar{r}_j \) tonnes/acre. b) That the planting cost is \( c_{1j} \) euros/acre. c) That \( b_{2j} \) tonnes are needed for next feed manufacturing season, for all \( j \in J = \{1, \ldots , J\} \). These amounts can be raised on the manufacturer land or bought from the market and any production in excess would be sold. For each crop \( j \), the buying and selling prices for the manufacturer are \( f_{2j} \) and \( g_{2j} \), respectively. In next table we summarize the parameters of this problem.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( j )</td>
<td>-</td>
<td>-</td>
<td>Index for crops</td>
</tr>
<tr>
<td>( J )</td>
<td>50</td>
<td>-</td>
<td>Number of crops</td>
</tr>
<tr>
<td>( J )</td>
<td>{1, \ldots , J}</td>
<td>-</td>
<td>Index set for crops</td>
</tr>
<tr>
<td>( c_{1j} )</td>
<td>( 100 + 3j )</td>
<td>euros / acre</td>
<td>Planting cost ( j \in J )</td>
</tr>
<tr>
<td>( \bar{r}_j )</td>
<td>( 4 + 0.02j )</td>
<td>tonnes / acre</td>
<td>Expected crop yield rate ( j \in J )</td>
</tr>
<tr>
<td>( b_1 )</td>
<td>1,700</td>
<td>acres</td>
<td>Available land to plant the crop</td>
</tr>
<tr>
<td>( b_{2j} )</td>
<td>( 100 + 3j )</td>
<td>tonnes</td>
<td>Amount of crop ( j ) needed ( j \in J )</td>
</tr>
<tr>
<td>( f_{2j} )</td>
<td>( 2c_{1j}/\bar{r}_j )</td>
<td>euros / tonne</td>
<td>Buying price of crop ( j ) ( j \in J )</td>
</tr>
<tr>
<td>( g_{2j} )</td>
<td>( 0.5 c_{1j}/\bar{r}_j )</td>
<td>euros / tonne</td>
<td>Selling prices of crop ( j ) ( j \in J )</td>
</tr>
</tbody>
</table>

It is clear that this is a synthetic example which will be used to illustrate the use of the Conditional Expectation (CE) approach as an improvement of the Expected Value (EV) counterpart.

The feed manufacturer wants to decide how much land to devote to each crop in order obtain the feed ingredients for the next season at the minimum cost. The first step to solve this problem is to define the decision variables:

<table>
<thead>
<tr>
<th>Decisions</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{x}_{1j} )</td>
<td>acres</td>
<td>Land devoted to crop ( j ) ( j \in J )</td>
</tr>
<tr>
<td>( \bar{y}_{2j} )</td>
<td>tonnes</td>
<td>Amount of crop ( j ) purchased (under production shortage) ( j \in J )</td>
</tr>
<tr>
<td>( \bar{z}_{2j} )</td>
<td>tonnes</td>
<td>Amount of crop ( j ) sold (under production excess) ( j \in J )</td>
</tr>
</tbody>
</table>

In order to decide how much land to devote to each crop we can formulate and solve the following EV problem:

\[
\min_x \quad z_{EV} = c_1^T \bar{x}_1 + f_2^T \bar{y}_2 - g_2^T \bar{z}_2 \tag{29}
\]

s.t. \( A_1 \bar{x}_1 \leq b_1 \) \( \tag{30} \)

\[
\bar{B}_2 \bar{x}_1 + \bar{y}_2 - \bar{z}_2 = b_2 \tag{31}
\]

\[
\bar{x}_1, \bar{y}_2, \bar{z}_2 \geq 0. \tag{32}
\]
where

\[ A_1 = \begin{pmatrix} 1, \ldots, 1 \end{pmatrix}_{1 \times J} \]

\[ \bar{B}_2 = \begin{pmatrix} \bar{r}_{2,1} & 0 \\ \vdots & \ddots \\ 0 & \bar{r}_{2,J} \end{pmatrix} \]

- In (29) the total cost is computed as the planting cost \( c_1^T \bar{x}_1 \) plus the crop purchasing cost \( f_2^T \bar{y}_2 \) (under production shortage) minus the crop selling revenue \( g_2^T \bar{z}_2 \) (under production excess).

- Constraint (30) states that the planted land \( A_1 \bar{x}_1 \) must be no more than the available land \( b_1 \).

- In (31) there is a balance equation such that crop production \( \bar{B}_2 \bar{x}_1 \) must be equal to crop demand \( b_2 \) with the help of purchasing or selling some crops, \( \bar{y}_2 \) or \( \bar{z}_2 \), respectively, if necessary.

After solving the previous EV problem by using the LP solver (CPLEX) one obtains:

\[ z_{EV}^* = 383,588 \]

\[ \bar{x}_1^* = \begin{pmatrix} 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 18.4 \ 31.0 \ 31.5 \ 32.1 \ 32.6 \ 33.1 \ 33.7 \ 34.3 \\ 34.8 \ 35.3 \ 35.8 \ 36.4 \ 36.9 \ 37.4 \ 37.9 \ 38.4 \\ 38.9 \ 39.4 \ 39.9 \ 40.4 \ 40.8 \ 41.3 \ 41.7 \ 42.2 \\ 42.7 \ 43.2 \ 43.6 \ 44.7 \ 44.5 \ 45.0 \ 45.4 \ 45.8 \\ 46.3 \ 46.7 \ 47.1 \ 47.5 \ 48.0 \ 48.4 \ 48.8 \ 49.2 \\ 49.6 \ 50.0 \ \end{pmatrix}^T \]

\[ \bar{y}_1^* = \begin{pmatrix} 103.0 \ 106.0 \ 109.0 \ 112.0 \ 115.0 \ 118.0 \ 121.0 \ 124.0 \\ 50.20 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \\ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 \ \end{pmatrix}^T \]

\[ \bar{z}_1^* = 0 \in \mathbb{R}^{+50}. \]

That is, devote the number of acres to each crop indicated by \( \bar{x}_1^* \) and purchase the amount of each crop indicated by \( \bar{y}_1^* \), to obtain the optimal EV cost (383,588 euros). Notice that the first eight feed ingredients will not be cultivated (they will be bought in the market). Please, be aware that we have broken vectors \( \bar{x}_1^* \) and \( \bar{y}_1^* \) into rows given their long lengths.

\[ \square \]

7.2 The CE feed manufacturer problem

Example 4. (The CE feed manufacturer problem: uncertain yields)

After solving the previous EV problem the feed manufacturer is a little worried since he knows that the yield rates are only expected values (the real values are uncertain). Let us assume that the uncertain yield rates can be modeled as a multinormal random vector \( \bar{r}_2 = (r_{2,1}, \ldots, r_{2,J})^T \). Now the problem
is a two-stage one: In the first stage the manufacturer decides the amount of land to devote to each crop. In the second stage the manufacturer will balance the need of each crop by buying or selling according to the yield observed. The description of the random yield rates and the related parameters can be found in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_2$</td>
<td>-</td>
<td>tonnes/acre</td>
<td>Random crop yield rates</td>
</tr>
<tr>
<td>$\mu_j$</td>
<td>$4 + 0.02j$</td>
<td>tonnes/acre</td>
<td>Expected yield rate of crop $j$</td>
</tr>
<tr>
<td>$\sigma_j$</td>
<td>$0.35 + 0.04j$</td>
<td>tonnes/acre</td>
<td>Standard deviation of $r_{2j}$</td>
</tr>
<tr>
<td>$\rho_{j_1,j_2}$</td>
<td>0.7</td>
<td></td>
<td>Correlation between $r_{2j_1}$ and $r_{2j_2}$ for $j_1 \neq j_2$ in $\mathcal{J}$</td>
</tr>
<tr>
<td>$\sigma_{j_1,j_2}$</td>
<td>$\rho_{j_1,j_2} \sigma_{j_1} \sigma_{j_2}$</td>
<td>(tonnes/acre)$^2$</td>
<td>Covariance between $r_{2j_1}$ and $r_{2j_2}$ for $j_1 \neq j_2$ in $\mathcal{J}$</td>
</tr>
<tr>
<td>$R$</td>
<td>50</td>
<td></td>
<td>Number of random parameters</td>
</tr>
<tr>
<td>$\mathcal{R}$</td>
<td>${1, \ldots, R}$</td>
<td>-</td>
<td>Index set for random parameters</td>
</tr>
<tr>
<td>$E_r$</td>
<td>33</td>
<td></td>
<td>Number of conditional scenarios derived from the $r$th conditional expectation</td>
</tr>
<tr>
<td>$\mathcal{E}_r$</td>
<td>${1, \ldots, E_r}$</td>
<td>-</td>
<td>Index set for conditional scenarios derived from the $r$th conditional expectation</td>
</tr>
<tr>
<td>$e$</td>
<td>-</td>
<td></td>
<td>Index for conditional scenarios</td>
</tr>
</tbody>
</table>

The feed manufacturer wants to decide how much land to devote to each crop in order to obtain the feed for the next season at the minimum expected cost. In the following table we have the decision variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{x}_{1j}$</td>
<td>acres</td>
<td>Land devoted to crop $j$</td>
</tr>
<tr>
<td>$\tilde{y}_{2j}^{re}$</td>
<td>tonnes</td>
<td>Amount of crop $j$ purchased under production shortage</td>
</tr>
<tr>
<td>$\tilde{z}_{2j}^{re}$</td>
<td>tonnes</td>
<td>Amount of crop $j$ sold under production excess</td>
</tr>
</tbody>
</table>

In order to decide how much land to devote to each crop we can formulate and solve the CE problem. As a first step we have calculated the conditional scenarios, that approximate the multivariate random vector $r_2$, by using Method 1. In total there are $R = 50$ random parameters and we fix the number of conditional scenarios per each random parameter $E_r = 33$ for all $r \in \mathcal{R}$. In total, we approximate $r_2$
by 1650 conditional scenarios. Then the CE problem (23)–(28), in this case, can be written as follows:

\[
\min_{\hat{x},\hat{y},\hat{z}} z_{CE} = c_1^T \hat{x}_1 + \frac{1}{R} \sum_{re \in R E} \hat{p}^r e (f_2^T \hat{y}^r e - g_2^T \hat{z}^r e)
\]

s.t. \[A_1 \hat{x}_1 \leq b_1 \]
\[
\hat{B}_{1e}^r \hat{x}_1 + \hat{y}_{2e} - \hat{z}_{2e} = b_2 \quad e \in E_1
\]
\[
\hat{B}_{2e}^r \hat{x}_1 + \hat{y}_{2e} - \hat{z}_{2e} = b_2 \quad e \in E_2
\]
\[
\vdots
\]
\[
\hat{B}_{Re}^r \hat{x}_1 + \hat{y}_{Re} - \hat{z}_{Re} = b_2 \quad e \in E_R
\]
\[
\hat{x}_1, \hat{y}_{2e}^r, \hat{z}_{2e}^r \geq 0 \quad re \in R E_r
\]

where

\[
\hat{B}_{2e}^r = \begin{pmatrix}
\hat{r}_{2,1}^{r e} & 0 \\
\vdots & \ddots \\
0 & \hat{r}_{2,R}^{r e}
\end{pmatrix}
\]

- In (33) we have the planting cost plus the expected cost to match the crop needs. In the case of good yield rates this expected cost could be negative, which would represent a profit.
- Constraint (34) states that the planted land must be no more than the available land.
- Constraints (35)–(37) are balance equations such that: First, the crops are planted according to \( \hat{x}_1 \). Second, after some weeks, the crop production \( \hat{B}_{2e}^r \hat{x}_1 \) is observed (assuming the conditional scenario \( re \)). Third, in order to match the crop needs, the crop production is augmented or reduced by purchasing \( \hat{y}_{2e}^r \) or selling \( \hat{z}_{2e}^r \), respectively, if necessary. In each row \( r \) we have the balance equations associated to the conditional scenarios derived from the \( r \)th conditional expectation.

After solving the previous CE problem by using the LP solver (CPLEX) one obtains:

\[
\begin{align*}
z_{CE}^* &= 420, 470 \\
\hat{x}_1^* &= \begin{pmatrix}
23.3 & 23.7 & 24.5 & 24.8 & 25.2 & 25.6 & 26.5 & 26.9 \\
27.3 & 27.6 & 28.0 & 28.3 & 28.7 & 29.0 & 30.3 & 30.7 \\
31.1 & 31.4 & 31.8 & 32.1 & 32.5 & 32.8 & 33.1 & 33.4 \\
33.8 & 34.1 & 34.4 & 34.7 & 35.0 & 35.3 & 35.6 & 36.1 \\
38.1 & 38.4 & 38.7 & 39.1 & 39.4 & 39.7 & 40.0 & 40.3 \\
40.6 & 40.9 & 41.2 & 41.5 & 41.8 & 42.1 & 42.3 & 42.6 \\
42.9 & 43.2 & & & & & & &
\end{pmatrix}^T.
\end{align*}
\]

That is, devote the number of acres indicated by \( \hat{x}_1^* \) to each crop to obtain the optimal CE cost (420,470 euros). As in the EV problem, we have broken vector \( \hat{x}_1^* \) into rows given his long length. Unlike in Example 3, here we do not report the optimal vectors \( \hat{y}_{2} \) and \( \hat{z}_{2} \) given their high dimensions.

### 7.3 Comparing the CE and the EV solutions

**Example 5.** *(The EEV and ECE feed manufacturer problems: random yields)*
With the data of Examples 3 and 4, let us compare the values EEV, ECE, CE and EV, which are the optimal values of the corresponding problems (EEV stands for the Expected result of using the EV solution and ECE, analogously). To compute the EEV we formulate the following EEV problem:

$$\min \tilde{x}, \tilde{y}, \tilde{z}$$

$$\text{EEV} = c^\top_1 \tilde{x} + \sum_{s \in S} \tilde{p}^s (f^\top_2 \tilde{y}^s - g^\top_2 \tilde{z}^s)$$

s.t. $A_1 \tilde{x} \leq b_1$

$$\tilde{B}^s_2 \tilde{x} + \tilde{y}^s - \tilde{z}^s = b^s_2, \quad s \in S$$

$$\tilde{x}_1, \tilde{y}^s, \tilde{z}^s \geq 0, \quad \tilde{x}_1 = x_{ref}$$

where we set the reference vector as $x_{ref} = \bar{x}_{EV}$. Notice that this is nothing but the RP problem with the additional constraint $\tilde{x}_1 = \bar{x}_{EV}$. It is well known [7] that the optimal value of this LP problem can be equivalently computed as

$$z^*_\text{EEV} = c^\top_1 x_{ref} + \sum_{s \in S} \tilde{p}^s \left( f^\top_2 [b_2 - \tilde{B}^s_2 x_{ref}]_+ - g^\top_2 [\tilde{B}^s_2 x_{ref} - b^s_2]_+ \right)$$

where $[\cdot]_+$ is the positive part, that is, given a vector, say $v$, then $[v]_+ = \max\{0, v\}$ componentwise.

To set our RP instance, we have used a set of $S = 10^7$ scenarios $\{\tilde{B}^s_2\}_{s \in S} \equiv \{\tilde{r}^s_2\}_{s \in S}$, sampled from the multinormal random vector $r_2 \sim N_{50}(\mu, \Sigma)$. In this case one has $\tilde{p}^s = 1/S$ for all $s \in S$.

<table>
<thead>
<tr>
<th>EV</th>
<th>EEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (euros)</td>
<td>383,588</td>
</tr>
<tr>
<td>Variation (%)</td>
<td>-</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1</td>
</tr>
</tbody>
</table>

In Table 1 we have the optimal EV and EEV costs. In the first row (label ‘Cost’), we observe EV ≤ EEV as predicted by Propositions 3 and 6. Notice that the EV problem ‘promises’ an optimal cost which is misleading, since the true expected costs, EEV, is worse than the ‘promised’ one. In the second row (label ‘Variation’) we have the variation relative to the EV value. Considering that in this problem, by Propositions 3 and 6, one has EV ≤ RP ≤ EEV, this variation can be interpreted as a worst-case optimality gap which can be computed as (EEV - EV)/EV = +18.4%. After considering these figures, a decision maker based on the EV approach could decide either to implement the EV optimal decision or to formulate and solve the RP problem based on scenarios, knowing that in the best case the improvement of the EEV would be 18.4%.

On the other hand, the ECE (Expected result of using the CE solution) can be obtained by solving the previous problem but setting the reference vector as $x_{ref} = \bar{x}_{CE}$ instead of $x_{ref} = \bar{x}_{EV}$. The results have been summarized in the following table (we also report the EV ones for comparison): In the first row, we observe EV ≤ CE ≤ ECE as predicted by Proposition 3. As in the EV problem, now the CE
problem ‘promises’ an optimal cost (420,470 euros) which is different from the true expected CE cost given by the ECE (437,227 euros). Now, in the second row (label ‘Variation’) we have the variation relative to the CE value, the tighter bound. Considering that in this problem \( CE \leq RP \leq (ECE\ or\ EVE) \), this variation can be interpreted as a worst-case optimality gap, which is \( (ECE - CE)/CE = 4.0\% \) and \( (EEV - CE)/CE = 8.0\% \) for problems CE and EV, respectively. The EV lower bound is 8.8\% worse than the EV counterpart. On the other hand, the ECE is 4\% better than the EEV terms of worst-case optimality gap. Finally, in the third row we observe that the total computing time has been 1+17 and 3+17 seconds for the EV and CE approaches, respectively. In summary, the CE approach, based on 1650 conditional scenarios, has outperformed the EV approach, based on 1 scenario (the expected one), at a moderate extra computational cost.

After considering these figures, a decision maker based on the CE approach could decide either to implement the CE optimal decision or to formulate and solve the RP problem based on scenarios, knowing that in the best case the improvement of the ECE would be 4\% (in contrast with the above 18.4\% obtained by the decision maker exclusively based on the EV method). The RP problem based on Cartesian scenarios would need at least \( 2^{30} \approx 1.13 \cdot 10^{15} \) scenarios. It is clear that the decision maker based on the RP approach would consider a much smaller number of scenarios to obtain a tractable RP problem by means, for example, of the moment matching method or the Sample Average Approximation (SAA) method, as already pointed out in previous sections.

8 Conclusions

In this paper we have considered the Two-stage Stochastic Linear Programming (TSLP) problem whose uncertain parameters are modeled as continuous random variables. The Recourse Problem (RP) and the Expected Value (EV) problem are two discrete approximations to the TSLP problem, based on scenarios and on the parameter expectations, respectively. The contributions of this paper have been: to introduce the Conditional Expectation (CE) problem, a new approximation to the TSLP problem based on conditional scenarios (a new concept here introduced) and to propose and analyze some useful bounds related to the CE problem. The message of this paper is twofold: a) The RP problem is a good approximation to the TSLP problem. b) However, if the computational effort to solve the RP problem results too high, our suggestion is to use the CE problem which requires a moderate computational effort and favorably compares to the EV problem in order to deal with parameter uncertainty.

From a practical point of view the CE approach is appealing since the computation of the conditional scenarios is straightforward. It only requires to compute and discretize a set of conditional expectations of the random vector that models the uncertain problem parameters. The number of conditional scenarios thus constructed, say \( N \), grows linearly with the number of random parameters of the TSLP problem. In this way the size of the resulting CE problem has a moderate size (\( N \) times the size of the EV problem), which can be solved with a moderate computational effort. Of course there is a price to be paid for this: the optimal CE solution is, in general, suboptimal for the RP problem but hopefully better than the optimal EV solution. Thus, for example, in the numerical example above solved the worst-case optimality gaps of the optimal CE and EV solutions have been 4\% and 8\%, respectively (Table 2).

From a theoretical point of view we have shown that in some cases the optimal RP cost can be bounded by the optimal CE and EV costs. In these cases, the EV bound is worse than the CE bound. Thus, for example, in the above numerical example the EV bound has resulted, at least, 8.8\% worse than the CE bound (Table 2). The reason is that the CE problem is a better approximation to the RP problem than the EV counterpart. Therefore, one would expect that the Expected result of using the EV solution (EEV) would be outperformed by the CE counterpart, which we have called the Expected result of using the CE solution (ECE). Although we have observed this result empirically, so far we have not a proof for this hypothesis and it remains a topic to be investigated.
As a matter of further research, apart from trying to prove the previous hypothesis, we are planning to analyze the use of conditional scenarios in other types of problems as for example mixed-integer linear programming and/or non linear programming. Furthermore, as a first step, we have applied the conditional scenario concept in a risk-neutral model. That is, it is based on the expected cost and it does not incorporate any risk measure as for example conditional value-at-risk [25] and stochastic dominance [12]. As a second step, we plan to incorporate some risk measure to our models based on conditional scenarios.

Acknowledgments: We are grateful to Jean-Philippe Vial and Alain Haurie for their comments and support at Logilab, University of Geneva, Switzerland. We also thank the support of the grant S2009/esp-1594 (Riesgos CM) from Comunidad de Madrid (Spain) and the grants MTM2012-36163-C06-06 from the Spanish Ministry of Economy and Competitiveness.

9 Appendix: Proofs of the theoretical results

In this section we prove the theoretical results of Section 6.

**Proposition 3.** Let us consider the RP problem (6)–(9) where some or all the components of $B_2$ and/or $b_2$ are stochastic (the other parameters being deterministic). Then

$$ z^*_{EV} \leq z^*_{CE} \leq z^*_{RP}. $$

**Proof.** We prove first the second inequality $z^*_{CE} \leq z^*_{RP}$. First, let us consider the RP problem with stochastic $B_2$ and $b_2$

$$ \min_x \quad z_{RP} = c_1^T \tilde{x}_1 + E[c_2^T \tilde{x}_2(\xi)] $$

s.t. $A_1 \tilde{x}_1 = b_1$

$$ \tilde{B}_2 \tilde{x}_1 + A_2 \tilde{x}_2(\xi) = \tilde{b}_2 $$

w.p.1

$$ \tilde{x}_2(\xi) \geq 0 $$

w.p.1

$$ \tilde{x}_1 \geq 0 $$

where $\tilde{\xi} = \text{vec}(\tilde{B}_2, \tilde{b}_2)$.

Second, let us define the RP2 problem as the RP problem with the second stage constrains (41) and (42) repeated $R$ times. Obviously the RP and RP2 problems are equivalent. Third, we obtain the RP3 problem by applying the conditional expectation operator $E[\cdot | \tilde{\xi}_r]$ for all $r \in \mathcal{R}$, to these $R$ sets of constraints (for details regarding the conditional expectation operator see [3]). The RP3 problem is as follows:

$$ \min_x \quad z_{RP3} = c_1^T \tilde{x}_1 + c_2 E[\tilde{x}_2(\tilde{\xi})] $$

s.t. $A_1 \tilde{x}_1 = b_1$

$$ E[\tilde{B}_2 | \tilde{\xi}_r] \tilde{x}_1 + A_2 E[\tilde{x}_2(\tilde{\xi}) | \tilde{\xi}_r] = E[\tilde{b}_2 | \tilde{\xi}_r] $$

w.p.1, $r \in \mathcal{R}$

$$ E[\tilde{x}_2(\tilde{\xi}) | \tilde{\xi}_r] \geq 0, $$

w.p.1, $r \in \mathcal{R}$

$$ \tilde{x}_1 \geq 0. $$

By the law of total expectation [8, 13], one has that

$$ E[\tilde{x}_2(\tilde{\xi})] = E[E[\tilde{x}_2(\tilde{\xi}) | \tilde{\xi}_r]] \quad \forall r \in \mathcal{R} $$

from where

$$ E[\tilde{x}_2(\tilde{\xi})] = \frac{1}{R} \sum_{r \in \mathcal{R}} E[E[\tilde{x}_2(\tilde{\xi}) | \tilde{\xi}_r]]. $$

18
Thus problem (44)–(48) can be written as

\[
\min_{\tilde{x}_1, \mathbb{E}[\tilde{x}_2(\xi) | \xi_r]} \quad z_{RP3} = c_1^T \tilde{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[c_2^T \mathbb{E}[\tilde{x}_2(\xi) | \xi_r]]
\]

s.t. \[ A_1 \tilde{x}_1 = b_1 \]
\[
\mathbb{E}[\tilde{B}_2 | \tilde{\xi}_r] \tilde{x}_1 + A_2 \mathbb{E}[\tilde{x}_2(\xi) | \tilde{\xi}_r] = \mathbb{E}[\tilde{b}_2 | \tilde{\xi}_r] \quad \text{w.p.1, } r \in R
\]
\[
\mathbb{E}[\tilde{x}_2(\xi) | \tilde{\xi}_r] \geq 0, \quad \text{w.p.1, } r \in R
\]
\[
\tilde{x}_1 \geq 0,
\]

which is not difficult to see that is equivalent to the following problem formulated in the \( \tilde{x} \) decision space.

\[
\min_{\tilde{x}_1} \quad c_1^T \tilde{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[c_2^T \tilde{x}_2(\xi_r)]
\]

s.t. \[ A_1 \tilde{x}_1 = b_1 \]
\[
\tilde{B}_2 \tilde{x}_1 + A_2 \tilde{x}_2(\xi_r) = \tilde{b}_2 \quad \text{w.p.1, } r \in R
\]
\[
\tilde{x}_2(\xi_r) \geq 0 \quad \text{w.p.1, } r \in R
\]
\[
\tilde{x}_1 \geq 0,
\]

where

\[
\tilde{x}_2(\xi_r) = \mathbb{E}[\tilde{x}_2(\xi) | \xi_r] \quad r \in R
\]
\[
\tilde{B}_2 = \mathbb{E}[\tilde{B}_2 | \tilde{\xi}_r] \quad r \in R
\]
\[
\tilde{b}_2 = \mathbb{E}[\tilde{b}_2 | \tilde{\xi}_r] \quad r \in R
\]
\[
\tilde{\xi}_r = \mathbb{E}[\tilde{\xi} | \tilde{\xi}_r] \quad r \in R.
\]

Notice that this model corresponds to the CE problem (16)–(22). Since it has been obtained by the constraint aggregation induced by the conditional expectation operator applied to the RP constrains, it is a relaxation of the RP problem and therefore \( z_{CE}^* \leq z_{RP}^* \).

Second, we prove the first inequality \( z_{EV}^* \leq z_{CE}^* \). Given the previous RP3 problem (44)–(48) we obtain the following RP4 problem by applying the expectation operator \( \mathbb{E}[\cdot] \) to constraints (46)-(47):

\[
\min_{\tilde{x}_1, \mathbb{E}[\tilde{x}_2(\xi) | \xi_r]} \quad z_{RP4} = c_1^T \tilde{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[c_2^T \mathbb{E}[\tilde{x}_2(\xi) | \tilde{\xi}_r]]
\]

s.t. \[ A_1 \tilde{x}_1 = b_1 \]
\[
\mathbb{E}[\mathbb{E}[\tilde{B}_2 | \tilde{\xi}_r]] \tilde{x}_1 + A_2 \mathbb{E}[\mathbb{E}[\tilde{x}_2(\xi) | \tilde{\xi}_r]] = \mathbb{E}[\mathbb{E}[\tilde{b}_2 | \tilde{\xi}_r]] \quad r \in R
\]
\[
\mathbb{E}[\mathbb{E}[\tilde{x}_2(\xi) | \tilde{\xi}_r]] \geq 0, \quad \text{w.p.1, } r \in R
\]
\[
\tilde{x}_1 \geq 0,
\]

which, by the law of total expectation, is equivalent to

\[
\min_{\tilde{x}_1, \mathbb{E}[\tilde{x}_2(\xi)]} \quad z_{RP4} = c_1^T \tilde{x}_1 + c_2^T \mathbb{E}[\tilde{x}_2(\xi)]
\]

s.t. \[ A_1 \tilde{x}_1 = b_1 \]
\[
\mathbb{E}[\tilde{B}_2] \tilde{x}_1 + A_2 \mathbb{E}[\tilde{x}_2(\xi)] = \mathbb{E}[\tilde{b}_2]
\]
\[
\mathbb{E}[\tilde{x}_2(\xi)] \geq 0
\]
\[
\tilde{x}_1 \geq 0.
\]
or equivalently

\[
\begin{align*}
\min_{\bar{x}_1, \bar{x}_2} \quad & c_1^T \bar{x}_1 + c_2^T \bar{x}_2 \\
\text{s.t.} \quad & A_1 \bar{x}_1 = b_1 \\
& \quad \bar{B}_2 \bar{x}_1 + A_2 \bar{x}_2 = \bar{b}_2 \\
& \quad \bar{x}_2 \geq 0 \\
& \quad \bar{x}_1 \geq 0.
\end{align*}
\]

where

\[
\begin{align*}
\bar{x}_2 &= \mathbb{E}[\bar{x}_2(\xi)] \\
\bar{B}_2 &= \mathbb{E}[\bar{B}_2] \\
\bar{b}_2 &= \mathbb{E}[\bar{b}_2]
\end{align*}
\]

Notice that this model corresponds to the EV problem (10)–(13). Since it has been obtained by the constraint aggregation induced by the expectation operator applied to the RP3 (equivalent to the CE problem), it is a relaxation of the CE problem and therefore \( z_{EV}^* \leq z_{CE}^* \).

**Proposition 4.** Let us consider the RP problem (6)–(9) where some or all the components of \( c_2 \) are stochastic (the other parameters being deterministic). Then

\[ z_{RP}^* \leq z_{CE}^* \leq z_{EV}^*. \]

**Proof.** Let us prove the first inequality \( z_{RP}^* \leq z_{CE}^* \). First we consider the RP problem with stochastic \( c_2 \)

\[
\begin{align*}
\min_{\bar{x}} \quad z_{RP} &= c_1^T \bar{x}_1 + \mathbb{E}[[c_2^T \bar{x}_2(\xi)]] \\
\text{s.t.} \quad & A_1 \bar{x}_1 = b_1 \\
& \quad \bar{B}_2 \bar{x}_1 + A_2 \bar{x}_2(\xi) = \bar{b}_2 \quad \text{w.p.}1 \quad (51) \\
& \quad \bar{x}_2(\xi) \geq 0 \quad \text{w.p.}1 \quad (52) \\
& \quad \bar{x}_1 \geq 0. \quad (53)
\end{align*}
\]

where \( \bar{\xi} = \bar{c}_2 \). Second, we define the RP5 problem as

\[
\begin{align*}
\min_{\bar{x}, \hat{x}} \quad z_{RP5} &= c_1^T \bar{x}_1 + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}[[\mathbb{E}[c_2^T \bar{x}_2(\xi) | \xi_r]]] \\
\text{s.t.} \quad & A_1 \bar{x}_1 = b_1 \\
& \quad B_2 \bar{x}_1 + A_2 \hat{x}_2(\xi) = b_2 \quad \text{w.p.}1, r \in \mathcal{R} \quad (56) \\
& \quad [\hat{x}_2(\xi) | \xi_r] \geq 0 \quad \text{w.p.}1, r \in \mathcal{R} \quad (57) \\
& \quad [\hat{x}_2(\xi) | \xi_r] = \hat{x}_2(\xi_r) \quad \text{w.p.}1, r \in \mathcal{R} \quad (58) \\
& \quad \bar{x}_1 \geq 0. \quad (59)
\end{align*}
\]

where (56)-(57) are equivalent to (51)-(52). On the other hand, by using the definition of \( \hat{x}_2(\xi_r) \) in equation (58), one has:

\[
\begin{align*}
\mathbb{E}[c_2^T \hat{x}_2(\xi) | \xi_r] &= \mathbb{E}[[c_2^T | \xi_r] \hat{x}_2(\xi) | \xi_r]] \\
&= \mathbb{E}[[c_2^T | \xi_r] \hat{x}_2(\xi_r)]
\end{align*}
\]
Therefore, one can write the RP5 problem as follows:

\[
\begin{align*}
\min_{\tilde{x}, \hat{x}} \quad & z_{RP5} = c_1^T \tilde{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[\mathbb{E}[c_2^T | \xi_r] \hat{x}_2(\xi_r)] \\
\text{s.t.} \quad & A_1 \tilde{x}_1 = b_1 \\
& B_2 \tilde{x}_1 + A_2 \hat{x}_2(\xi_r) = b_2 \quad \text{w.p.1, } r \in \mathcal{R} \\
& [\tilde{x}_2(\xi) | \xi_r] = \hat{x}_2(\xi_r) \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_2(\xi_r) \geq 0 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_1 \geq 0.
\end{align*}
\]

which is not difficult to see that is equivalent to the following problem formulated in the \( \hat{x} \) decision space.

\[
\begin{align*}
\min_{\hat{x}} \quad & c_1^T \hat{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[\hat{c}_2^T \hat{x}_2(\xi_r)] \\
\text{s.t.} \quad & A_1 \hat{x}_1 = b_1 \\
& B_2 \hat{x}_1 + A_2 \hat{x}_2(\xi_r) = b_2 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_2(\xi_r) \geq 0 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_1 \geq 0.
\end{align*}
\]

where

\[
\hat{c}_2^T = \mathbb{E}[c_2^T | \xi_r] \quad r \in \mathcal{R}
\]

We observe that this model corresponds to the CE problem. It has been derived by including the additional constraints (58) in the RP problem, therefore it reduces the RP feasibility set and, as a consequence, \( z_{RP}^* \leq z_{CE}^* \).

To prove the second inequality \( z_{CE}^* \leq z_{EV}^* \), we can proceed in an analogous way and define the CE2 problem as follows

\[
\begin{align*}
\min_{\hat{x}} \quad & z_{CE2} = c_1^T \hat{x}_1 + \frac{1}{R} \sum_{r \in R} \mathbb{E}[\hat{c}_2^T \hat{x}_2(\xi_r)] \\
\text{s.t.} \quad & A_1 \hat{x}_1 = b_1 \\
& B_2 \hat{x}_1 + A_2 \hat{x}_2(\xi_r) = b_2 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_2(\xi_r) \geq 0 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_2(\xi_r) = \bar{x}_2 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_1 \geq 0.
\end{align*}
\]

which is equivalent to

\[
\begin{align*}
\min_{\hat{x}} \quad & z_{CE2} = c_1^T \hat{x}_1 + \left( \frac{1}{R} \sum_{r \in R} \mathbb{E}[\hat{c}_2^T] \right) \bar{x}_2 \\
\text{s.t.} \quad & A_1 \hat{x}_1 = b_1 \\
& B_2 \hat{x}_1 + A_2 \bar{x}_2 = b_2 \\
& \bar{x}_2 \geq 0 \\
& \hat{x}_2(\xi_r) = \bar{x}_2 \quad \text{w.p.1, } r \in \mathcal{R} \\
& \hat{x}_1 \geq 0.
\end{align*}
\]

Now, taking into account that

\[
\frac{1}{R} \sum_{r \in R} \mathbb{E}[\hat{c}_2] = \frac{1}{R} \sum_{r \in R} \mathbb{E}[\mathbb{E}[\hat{c}_2 | \xi_r]] = \frac{1}{R} \mathbb{E}[\hat{c}_2] = \bar{c}.
\]
problem CE2 is equivalent to the following problem formulated in the $\bar{x}$ space:

$$\min_{\bar{x}} \quad c_1^T \bar{x}_1 + c_2^T \bar{x}_2$$

s.t.  
$$A_1 \bar{x}_1 = b_1$$
$$B_2 \bar{x}_1 + A_2 \bar{x}_2 = b_2$$
$$\bar{x}_2 \geq 0$$
$$\bar{x}_1 \geq 0.$$  

We observe that this model corresponds to the EV problem. Since, it has been derived by including the additional constraints (64) to the CE problem, the CE feasibility set is reduced and, as a consequence, $z_{CE}^* \leq z_{EV}^*$.

**Proposition 5.** Let us consider the RP problem (6)–(9) where some or all the components of $c_2$ are stochastic (the other parameters being deterministic). Then

$$z_{RP}^* \leq z_{EV}^* \leq z_{EV}^* \quad \text{and} \quad z_{RP}^* \leq z_{ECE}^* \leq z_{CE}^*.$$  

**Proof.** Let us prove the right-hand side chain of inequalities. In the proof of Proposition 4, we saw that the CE problem is equivalent to the RP5 problem (54)–(59), which corresponds to the RP problem, with a set of additional constraints. In order to define the RP6 problem we add constraint (71) to the RP5 problem as follows:

$$\min_{\bar{x}} \quad z_{RP6} = c_1^T \bar{x}_1 + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}\left[\mathbb{E}\left[ c_2^T \bar{x}_2(\xi) \mid \xi_r \right] \right]$$

s.t.  
$$A_1 \bar{x}_1 = b_1$$
$$B_2 \bar{x}_1 + A_2 \bar{x}_2 = b_2 \quad \text{w.p.1, } r \in \mathcal{R}$$
$$\bar{x}_2(\xi) \mid \xi_r \geq 0 \quad \text{w.p.1, } r \in \mathcal{R}$$
$$\bar{x}_1 = \hat{x}_1^*$$
$$\bar{x}_1 \geq 0.$$  

It is obvious that $z_{RP6}^* = z_{RP5}^* = z_{CE}^*$. On the other hand, we can relax constraints (70) in the RP6 problem in order to obtain the ECE problem. Since the ECE problem is a relaxation of the RP6 problem we have $z_{ECE}^* \leq z_{RP6}^*$. Putting all together we have $z_{ECE}^* \leq z_{RP6}^* = z_{RP5}^* = z_{CE}^*$, which implies $z_{ECE}^* \leq z_{CE}^*$, as we wanted to prove.

Let us prove the left-hand side chain of inequalities. From the proof of Proposition 4, it can be seen that the EV problem is equivalent to the RP5 problem (54)–(59) plus constraint (64), that is, $\check{x}_2(\xi_r) = \check{x}_2$. Otherwise said, the EV problem corresponds to the RP problem with two sets of additional constraints. In order to define the RP7 problem we add constraints (78) and (79) to the RP5 problem as follows:

$$\min_{\check{x}} \quad z_{RP7} = c_1^T \check{x}_1 + \frac{1}{R} \sum_{r \in \mathcal{R}} \mathbb{E}\left[\mathbb{E}\left[ c_2^T \check{x}_2(\xi) \mid \xi_r \right] \right]$$

s.t.  
$$A_1 \check{x}_1 = b_1$$
$$B_2 \check{x}_1 + A_2 [\check{x}_2(\xi) \mid \xi_r] = b_2 \quad \text{w.p.1, } r \in \mathcal{R}$$
$$\check{x}_2(\xi) \mid \xi_r \geq 0 \quad \text{w.p.1, } r \in \mathcal{R}$$
$$\check{x}_2(\xi_r) = \check{x}_2 \quad \text{w.p.1, } r \in \mathcal{R}$$
$$\check{x}_1 = \check{x}_1^*$$
$$\check{x}_1 \geq 0.$$  

22
It is clear that $z^*_\text{RP7} = z^*_\text{EV}$. On the other hand, we can relax constraints (77)–(78) in the RP7 problem in order to obtain the EEV problem. Since the EEV problem is a relaxation of the RP7 problem we have $z^*_\text{EEV} \leq z^*_\text{RP7}$. Putting all together we have $z^*_\text{EEV} \leq z^*_\text{RP7} = z^*_\text{CE}$, which implies $z^*_\text{EEV} \leq z^*_\text{EV}$, as we wanted to prove.

The previous bounds can only be applied when one restricts the set of stochastic parameters to $B_2, b_2$ or to $c_2$. Next result applies for all the cases.

**Proposition 6.** Let us consider the RP problem (6)–(9) where some or all the second stage parameters are stochastic ($c_1, A_1$ and $b_1$ are always supposed to be deterministic). Then

$$z^*_\text{RP} \leq z^*_\text{EEV} \quad \text{and} \quad z^*_\text{RP} \leq z^*_\text{ECE}.$$  

**Proof.** Let us prove the left-hand side inequality first. The reason is that to obtain $z^*_\text{EEV}$ we have to solve the RP problem with the additional constraint $\tilde{x}_1 = \bar{x}_1^*$, which tightens it. The same reason applies for the ECE case. □

**References**


