TWO-STAGE STOCHASTI LINEAR PROGRAMMING – SPLITTING METHOD

Luis Ernesto Torres Guardia
Universidade Federal Fluminense
Rua Passo da Pátria 156, São Domingos, Niterói, RJ, 24210-240
tepletg@vm.uff.br

Tania Guillén de Torres
Universidade Federal do Rio de Janeiro
Av. Horácio Macedo s/n, Ilha do Fundão – RJ, 21941-598
tguillen@iesc.ufrj.br

RESUMO

Neste trabalho, é apresentado o problema linear estocástico de dois estágios o qual é formulado como um problema de programação linear, representado por cenários, usando o método divisão. Uma técnica apropriada, como o método de pontos interiores, pode ser usada para resolver o programa linear explorando a estrutura da nova formulação para grande porte. O esforço computacional na maioria dos método de pontos interiores é em encontrar a solução do correspondente sistema linear de equações, e para isso, é usado a decomposição de Cholesky implementando no MATLAB. Alguns resultados numéricos preliminares são reportados.

PALAVARAS CHAVE. Programação linear estocástica, Método de pontos interiores, Decomposição de Cholesky.

ABSTRACT

In this work, it is presented the two-stage stochastic linear programming which is formulated as a linear programming problem, represented by scenarios, using the splitting method. An appropriate technique, such as interior point methods, can be used to solve the linear program exploiting the structure of the new formulation to a large extent. The computational effort in most interior point methods is dominated of finding the solution of the corresponding linear system of equations, and for this, it is used the Cholesky factorization in the MATLAB code. Some preliminary numerical results are reported.

KEYWORD. Stochastic linear programs. Interior point methods. Cholesky factorizations.

1.Introduction

Many practical problems with data under uncertain can be modeled as mathematical programs, and this case resulted in the field of stochastic programming and for the last years this theory of stochastic programming has attracted the attention of researchers and plays an increasingly important role in many situations in real-world applications. Some examples of this area of stochastic programming include portfolio management models, electric power generation capacity planning models, forestry management problems, as mentioned by Birge and Holmes (1992). Another example is a production planning problem as mentioned by Zhao (2001) where the demand is unknown when the production is planned and there is no way to assume that the amount produced can meet the demand exactly.

Most basic among stochastic programs are discretely stochastic linear programs, which are the extensions of any standard linear programs, and may lead to large stochastic linear programs and extensive computational efforts may be required since the size of these problems generally grows exponentially with the number of stochastic parameters in the formulation.

There are various ways how the stochastic problem can be modeled and many algorithms for solving this problem. The main difficulty is caused by its nature of the dimensionality. Interested readers are referred to the following books on stochastic programming, Birge and Louveaux (1997) and Kall and Wallace (1994), Alonso-Ayuso, et al. (2009), Infanger (1994), Shapiro et al. (2009) and related references therein.

As a result of the wide variety of applications, as mentioned above, a number of different stochastic optimization models appeared in the literature. In this work, it is concentrated on the two-stage stochastic program with finite discrete distribution on the random entries, that is, where uncertainty is incorporated into the problem by the use of scenarios. The resulting stochastic linear programming is dramatically large and requires very high performance in order to solve real world problems.

With the rapid growth, initiated by Karmarkar in 1984, and development in interior point methods to solve large linear programming problems, this method could be a successful solution method for stochastic programming and should exploit the special structure of the problem to cut down computational times.

The paper is organized as follows. In section 2, the two-stage stochastic linear programming is presented with a finite number of possible realizations, called scenarios. Section 3 describes the primal-dual interior-point to solve the corresponding linear programming. Section 4 introduces a key technique, a full-splitting method, to be employed in the linear problem. In section 5, it is presented some numerical results, via different number of scenarios, for the approach mentioned above. As an application, only the right-hand side of the linear programming is stochastic. Finally, the paper concludes in section 6.

2. Two-stage stochastic linear programming

Consider the following situation. There are two phases in a decision-making process. At the beginning of the first phase, one has to make a decision without precise knowledge of the random parameters in the second stage. In this phase, the structural component is fixed and free of any uncertainty. After observing uncertainty, it is corrected the decision of stage 1.

The typical two-stage stochastic program can be written in standard form as follows:

$$\begin{aligned} \text{Min} \quad & c^T x + E(\rho(x,w)) \\ \text{subject to:} \quad & Ax = b, \\ & x \geq 0, \end{aligned} \qquad (1)$$

where E stands for expectation with respect to the random variable w and ρ is the correction cost, also called the recourse function, defined as:

$$\rho(x,w) = \min\ d(w)^T\ y(w) \quad (2)$$
$$\text{subject to: } B(w)x + W(w)y(w) = h(w),$$
$$y(w) \geq 0.$$

for all possible realizations (or scenarios) w.
Problem (1) treats to minimize the sum of the first stage objective and the expected value of the recourse function.
Usually, x and y are referred to as first and second stage decision variables, respectively, and W(w) is called the recourse matrix. The decision variable $x \in R^n$ is a vector of first-stage variable that are subject to fixed, structural constraints, associated to the first stage constraint matrix A, and represent decisions that must be made before the values of uncertain parameters are observed. The first-stage decisions variables are often referred to as here-and-now decisions solutions. The values of these variables should be independent of the realization of uncertain parameters. The variable $y(w) \in R^s$ is the vector of second stage control decisions that represent recourse actions that can be taken after a specific realization of the uncertain parameters is observed. These second-stage decisions are referred as wait-and-see decisions solutions. The vectors $b \in R^m$ is the first stage right-hand side vector and $c \in R^n$ the corresponding cost vector. The matrix $A \in R^{mxn}$ and the vectors b and c are deterministic and $B(w) \in R^{lxn}$, $W(w) \in R^{lxv}$ are random matrices, and $h(w) \in R^l$ the second stage right-hand side vector, $d(w) \in R^v$ is the second stage cost vector.
A standard approach to solving the two stage problem (1) – (2) is by constructing scenarios, it means that the random data has a finitely supported distribution. That is, one generates a finite numbers of possible scenarios w_k which might occur and assigns to each w_k a probability $p_k > 0$, such that

$$\sum_{k=1}^{K} p_k = 1$$

where K is the number of scenarios. For example, the scenarios may represent historical data collected over a period of time.
Taking into account every scenario, the two-stage problem (1)-(2) can be formulated as the following deterministic and large equivalent optimization linear problem:

$$\begin{aligned} \text{Min } & c^Tx + p_1d_1^Ty_1 + \ldots + p_Kd_K^Ty_K && (3)\\ \text{subject to: } & Ax = b,\\ & B_1x + W_1y_1 = h_1,\\ & \quad\vdots\\ & B_Kx + W_Ky_K = h_K,\\ & x \geq 0,\ y_1 \geq 0,\ \ldots,\ y_K \geq 0. \end{aligned}$$

The first constraint of the above linear programming (3) represents m structural constraints whose coefficients are fixed and unaffected by uncertainty. The coefficient matrices B_v , W_v and the vectors h_v and d_v may take different values under each scenario.
The linear programming problem (3) has a special structure, the so called dual block- angular structure. For the above problem (3), the dual problem is given by:

$$\begin{aligned} \text{Max } & b^Tu + h_1^Tv_1 + \ldots + h_K^Tv_K && (4)\\ \text{subject to: } & A^Tu + B_1^Tv_1 + \ldots + B_K^Tv_K + s = c,\\ & W_1^Tv_1 + z_1 = p_1d_1,\\ & \quad\vdots\\ & W_K^Tv_K + z_K = p_Kd_K, \end{aligned}$$

$$s \geq 0, \;\; z_1 \geq 0, \;\; \ldots, \; z_K \geq 0.$$

It can be seen that the size of the equivalent linear programming problem (3) grows rapidly with the number of realizations, and in many applications the number of scenarios K may be very large, but sparse linear problem where the exploitation of the structure is very important. Some different special classes of the problem (4) are given in the literature of stochastic problem. If W is not a random, that is, $W_1 = W_2 = \ldots = W_K$, then the problem has fixed recourse. If the second-stage of a fixed recourse problem is feasible for every possible realization of the data random, it is said that the problem has a complete recourse. A special case of complete recourse is a simple recourse where W = [I, -I].
There are many algorithms for solving the two-stage problem (3). As mentioned by Zhao (2001), these algorithms can be roughly classified into three classes: (i) direct methods which directly use the simplex method or interior point methods to solve a linear program; (ii) cutting plane based decomposition methods which generate a set of cuts to approximate the nonlinear and non-smooth problem given in (1) and (iii) derivate based decomposition methods which determine a search direction by using the gradient and Hessian at the current point and find a new point along this direction. Interested readers can also see these methods in details from the papers written by Berkelaar, et al. (2002), Kall and Mayer (2006), Linderoth et al. (2006), Mészaros (1997), Sun and Liu (2006), Ruszczynski (1999), Ruszczynski, and Swietanowski, (1997), etc. Certainly, the selection of methods to be used depends on the two-stage problem to be solved.
This work is concentrated on methods that use specific stochastic program structure, for this, it is used the primal-dual interior-point method to solve the linear programming given in (3), using the full-splitting technique. The next section will describe the interior point method.

3. Primal-Dual Interior-Point Method

In practice, experience of extensive numerical tests indicates that the primal-dual interior-point method is the most powerful algorithm from a family of interior point methods.
Consider the following primal linear programming problem:

$$\text{(P)} \qquad \begin{aligned} \min \;\; & c^T x \\ \text{subject to:} \;\; & Ax = b, \\ & x \geq 0, \end{aligned} \tag{5}$$

where A is a m x n matrix, c and x are n vectors, and b is a m vector. It is assumed that A is of full row rank.
The dual problem associated with the primal linear programming (5) can be written as follows:

$$\text{(D)} \qquad \begin{aligned} \max \;\; & b^T y \\ \text{subject to:} \;\; & A^T y + z = c, \\ & z \geq 0, \end{aligned} \tag{6}$$

where y denote the m vector of dual variables and z is the n vector of dual slack variables.
Finding the optimal solutions of (P) and (D) is equivalent to solving the following system:

$$\begin{aligned} Ax = b, \quad & x \geq 0, \\ A^T y + z = c, \quad & z \geq 0, \\ XZe = 0, \quad & \end{aligned} \tag{7}$$

where X is a diagonal matrix, its components is given by the components of vector x, that is, $X = \text{diag}(x_1,\ldots,x_n)$, and Z is a diagonal matrix given by, $Z = \text{diag}(z_1,\ldots,z_n)$, and e is the all one vector, that is, $e = (1,1,\ldots,1)^T \in R^n$.

The basic idea of the primal-dual interior-point is to replace the third equation in (7) by the parameterized equation $XZe = \mu e$, where μ is a barrier parameter. This leads to the following system:

$$\begin{aligned} Ax = b,\ \ x \geq 0,\\ A^T y + z = c,\ \ z \geq 0,\\ XZe = \mu e. \end{aligned} \qquad (8)$$

Under certain conditions, it can be shown that as μ goes to zero in (8), the third equation in (7) is satisfied and it yields optimal solutions for both primal and dual problems.
The Newton's method is used to solve the system in (8). The resulting linear system is denominated the Newton equation system whose solutions (dx , dy, dz) are given by the following equations, denominated normal equations:

$$\begin{aligned} (A Z^{-1} X A^T) dy &= A Z^{-1} X(r_c - X^{-1} r_\mu) + r_b\\ dz &= - A^T dy + r_c .\\ dx &= Z^{-1} (r_\mu - X\, dz), \end{aligned} \qquad (9)$$

where:

$$\begin{aligned} r_c &= c - A^T y - z,\\ r_b &= b - Ax ,\\ r_\mu &= \mu e - XZe. \end{aligned}$$

To summarize an iteration of the primal-dual interior-point method, let, at the j-th iteration, $dw_j = (dx_j, dy_j, dz_j)^T$ denote the solution obtained for system (9). In the next iteration, a new point $w_{j+1} = (x_{j+1}, y_{j+1}, z_{j+1})^T$ is determined by

$$\begin{aligned} x_{j+1} &= x_j + \alpha_j dx_j\\ y_{j+1} &= y_j + \alpha_j dy_j\\ z_{j+1} &= z_j + \alpha_j dz_j, \end{aligned}$$

α_j being the step length, determined by a suitable line search procedure.
With this new point w_{j+1}, the barrier parameter μ is updated according to certain rules and a new system (9) is formed. It is solved by any solution method and the iterative procedure follows until a stopping rule is satisfied.

4.Full - Splitting Technique.

In this work, it is considered the full-splitting formulation, used with different format by Birge (1997), for problem (3) and it is represented in the following case:

$$\begin{array}{llr} \text{Min } c^T x + 0x_1 + \ldots + 0x_K + p_1 d_1^T y_1 + \ldots + p_K d_K^T y_K & & (10)\\ \text{subject to } Ax & = b, & \\ \quad B_1 x_1 + W_1 y_1 & = h_1, & \\ \quad \cdot \qquad \cdot & \cdot & \\ \qquad \cdot \qquad \cdot & \cdot & \\ \qquad\qquad B_K x_K + W_K y_K & = h_K, & \\ x - x_1 & = 0, & \\ \cdot \qquad \cdot & & \\ \cdot \qquad\qquad \cdot & & \\ x \qquad\qquad - x_K & = 0, & \\ x \geq 0, x_1 \geq 0, \ldots, x_K \geq 0, y_1 \geq 0, \ldots, y_K \geq 0, & & \end{array}$$

or it is convenient to write the problem (10) in the following compact form:

$$\begin{aligned} \min\ & c^T x + 0^T x_1 + \ldots + 0^T x_K + p_1 d_1^T y_1 + p_2 d_2^T y_2 + \ldots + p_K d_K^T y_K \\ \text{subject to}\ & A' x' = b', \\ & x' \geq 0, \end{aligned} \tag{11}$$

where the full constraint matrix has the form:

$$A' = \begin{bmatrix} A & & & & & \\ & B_1 & & W_1 & & \\ & & \ddots & & \ddots & \\ & & B_K & & & W_K \\ I & -I_1 & & & & \\ \vdots & & \ddots & & & \\ I & & -I_K & & & \end{bmatrix},$$

and $x' = (x, x_1, \ldots, x_K, y_1, \ldots, y_K)^T$, $c' = (c, 0, \ldots, 0, p_1 d_1, \ldots, p_K d_K)^T$ and $b' = (b_1, h_1, \ldots, h_K, 0 \ldots, 0)^T$ considering the appropriate dimensions.
To apply the primal-dual interior point method, the matrix $(A')(D')(A')^T$ is formed and in this case is given by:

$(A')(D')(A')^T =$

$$\begin{bmatrix} ADA^T & & & AD & . & AD \\ & B_1 D_1 B_1^T + W_1 V_1 W_1^T & & -B_1 D_1 & & \\ & & B_K D_K B_K^T + W_K V_K W_K^T & & & -B_K D_K \\ DA^T & -D_1 B^T & & D + D_1 & & D \\ . & & . & . & & \\ DA^T & & -D_K B^T & D & & D + D_K \end{bmatrix}$$

or simply in compact form:

$$(A')(D')(A')^T = \begin{bmatrix} F & O & J \\ O & C & K \\ J^T & K^T & E \end{bmatrix}$$

where $F = ADA^T$, $J = [\,AD \ldots AD\,]$, K is a diagonal matrix of block, $K = \text{diag}(-B_i D_i)$, $i = 1,\ldots,K$, C also a block diagonal matrix, $C = \text{diag}(B_i D_i B_i^T + W_i V_i W_i^T)$, $i = 1,\ldots,K$, and

$$E = \begin{bmatrix} D+D_1 & D & & D \\ D & D+D_2 & & \\ & & & \\ & & & D \\ D & & D & D+D_K \end{bmatrix},$$

where D ε R^n is a diagonal matrix, D = Z^{-1} X, D_i ε R^n a diagonal matrix, $D_i = (Z_i^{-1}) X_i$ and V_i ε R^v a diagonal matrix, $V_i = (Z_i')^{-1} Y_i$, for i = 1,…,K.
Following the interior point method, it is solved the following linear system:

$$(A')D(A')^T dy = r , \qquad (12)$$

where dy = $(dy_1, dy_2, dy_3)^T$ and r = $(r_1, r_2, r_3)^T$.
The solution of the above linear system (12) is given by:

$$[E - J^T F^{-1} J - K^T C^{-1} K] dy_3 = r_3 - J^T F^{-1} r_1 - K^T C^{-1} r_2, \qquad (13)$$
$$dy_1 = F^{-1} (r_1 - J dy_3), \qquad (14)$$
$$dy_2 = C^{-1} (r_2 - K dy_3), \qquad (15)$$

and

$$r = EZ^{-1}X\xi_{c'} - EZ^{-1}\xi_\mu + \xi_{b'},$$
$$\xi_{b'} = b' - A'x,$$
$$\xi_{c'} = c' - A'^T y - z,$$
$$\xi_\mu = \mu e - Xze.$$

The major problem, as any interior point method, is to solve the linear system given in (13). This linear system of equations requires determining the inverse of the block diagonal matrix C= diag($B_iD_iB_i^T + W_i V_iW_i^T$), i = 1,…,K,. In practice it is found the Cholesky factorization, or any other decomposition technique, of each matrix diag($B_iD_iB_i^T + W_iV_iW_i^T$), for each scenario i, i = 1,…,K, and use them to form the right hand side and the matrix associated to equation given in (13), and then find the related factorization of matrix [$E - J^T F^{-1} J - K^T C^{-1} K$].
For example, factor the block diagonal matrix C = diag($B_iD_iB_i^T + W_iV_iW_i^T$), i = 1,…,K, into $L_CL_C^T$ using Cholesky factorization, that is, C = $L_CL_C^T$ where L_C is lower triangular matrix. This operation can be implemented in parallel machines for each scenario. The same, using Cholesky factorization, factor the matrix [$E - J^T F^{-1} J - K^T C^{-1} K$] and solve the corresponding linear system. Instead of factor this matrix, Castro (2007) uses the preconditioned conjugate gradient method to solve the above normal equations. To find the inverse of matrix F, is not a difficult task because of its low dimension, factor matrix F using Cholesky factorization and apply to equation (13) and (14).

5.Computational Experiments

In this section, it is evaluated the performance of the splitting method for the stochastic problem developed in the last section. The set of test instances is generated based on a production/distribution problem, where the demand of the customers is unknown when the production on the set of all suppliers is planned. The kind of problem is studied by Cheung and Powell (1996) for the two-stage and multistage stochastic programming.
In all these cases of the production/distribution problems, it is not take the advantage the structure of the restrictions of these problems, that is, it is storaged all the matrices associated to the above stochastic problem. The exploting of these restrictions could be done in future work for large scale problems.

The full-splitting technique is programmed in MATLAB code and the test is conducted on a Core I5 laptop computer with 4GB of RAM**,** running at 2.53 GHz.
For simplicity, it is supposed that all scenarios occur with the same probability. The initial data of all the test problems is infeasible and given by x = 10, z = 1 and y = -0.25.
For each test problem, the demand h_k, the right-hand-term, is generated such that, $h_k' = d_k + 0.1$, where d_k is a fixed vector. Also, matrix W is not random, i.e., $W_1 = W_2 = \ldots = W_K$, the related problem has fixed recourse. Additionally, for the generated right hand side above, the second stage of the fixed recourse problem, it is said that the problem has a complete recourse.
The numerical results are presented in the following table, where μ is the barrier parameter. Iter represents the number of iterations. CPU represents the Central Processing Unit time (in seconds) for running the MATLAB code in solving the production/distribution problem, with 5 production centers, 7 warehouse centers and 10 customer centers. The number of scenarios, represented by nc, varies from 30 to 105.
The algorithm terminates when the following relative gap, computed by the formula:

$$\frac{(c')^T x' - (b')^T y'}{1.0 + \left|(b')^T y'\right|}$$

is less than 10^{-8}, where $(c')^T x'$ is the value of the objective function of the primal linear problem and $(b')^T y'$ is the value of the objective function of the corresponding dual problem.
The number of iterations in the interior point method is typically very low as it can be seen from the table (iter). Also, it can be noticed that the value of μ is very low. The time was measured with the cputime function of MATLAB and is recorded in the last line of the table.

Table. Results on Production/distribution Problems

nc	30	60	90	105
iter	7	7	7	7
μ	$3.0112e^{-9}$	$3.4079e^{-9}$	$3.7192e^{-9}$	$3.8769e^{-9}$
CPU	3.363898	20.483931	108.651003	614.685508

6. Conclusions

In this paper, it is proposed the splitting method to solve the stochastic linear programming formulated via scenarios. Generally, this kind of problem grows rapidly as the number of scenarios grows. For this, it is used the dual-dual interior-point method to solve the new large scale linear program. The method proposed in this work is intended to provide an alternative to others papers. All the matrices associated to the linear program are storaged without using any technique of exploiting the structure of these matrices, but the preliminary numerical experiments demonstrate encouraging performance of the proposed method. Further, the code shall be improved, can be done in parallel, and conduct more numerical experiments by exploiting the special structure, in this case, of the production/distribution problem. This is merely for illustrative purposes, although this decomposition method could be done for any problem. This will be a subject of future research.

References

Alonso-Ayuso, A., Escudero, L. and Pizarro-Romero, C., (2009), Introduction to Stochastic Programming. Dykinson, Madrid.

Berkelaar, A., Dert, C., Oldenkamp, B. and Zhang, S., (2002), A primal-dual decomposition-based interior point approach to two-stage stochastic linear programming. Operations Research, 50, 904 – 915.

Birge, J. and Holmes, D., (1992), Efficient solution of two-stage stochastic linear programs using interior point methods, Computational Optimization and Applications, 1, 245 – 276.
Birge, J. R., (1997), Stochastic programming computation and applications. INFORMS Journal on Computing, 9, 111 – 133.
Birge, J.R. and Louveaux, F.V. , Introduction to Stochastic Programming. Springer-Verlag, New York, 1997.
Castro, J., (2007), An interior-point approach for primal-angular problems, Computational Optimization and Applications, 36, 195 – 219.
Cheung, R. and Powell, W., (1996), Models and algorithms for distribution problems with uncertain demands. Transportation Science, 30, 43 – 59.
Infanger, G., Planning under uncertainty: Solving large-scale stochastic linear programs. Boyd and Fraser Publishing Company, 1994.
Kall, P. and Wallace, S.W., Stochasting Programming. John Wiley and Sons, 1994.
Kall, P. and Mayer, A., (2006), Some insights into the solution algorithms for SLP problems, Annals of Operations Research, 142, 147 – 164.
Linderoth, J. Shapiro, A. and Wright, S., (2006), The empirical behavior of sampling methods for stochastic programming. Annals of Operations Research, 142, 215-241.
Karmarkar, N., (1984), A polynomial-time algorithm for linear programming, Combinatorica, 4, 373 - 395.
Mészáros, C., (1997), The augmented system variant of IPMs in two-stage stochastic linear programming computation. European Journal of Operational Research, 101, 317 – 327.
Shapiro, A., Dentcheva, D., and Ruszczynski, A. Lectures on Stochastic Programming: Modeling and Theory, SIAM Philadelphia, 2009.
Sun, J. and Liu, X., (2006), Scenario formulation of stochastic linear programs and the homogeneous self-dual interior point method, Informs Journal on Computing, 18, 4, 444 - 454.
Ruszczynski, A., (1999), Some advances in decomposition methods for stochastic linear programming, Annals of Operations Research, 85, 153 – 172.
Ruszczynski, A. and Swietanowski, A., (1997), Accelerating the regularized decomposition method for two stage stochastic linear problems, Europen Journal of Operatinal Research, 101, 328 – 342.
Zhao, G., (2001), A log-barrier method with Benders decomposition for solving two-stage stochastic linear programs. Mathematical Programming, Series A, 90, 507-536.