A PRIMAL-DUAL ALGORITHM FOR THE MULTISTAGE STOCHASTIC DISTRIBUTION PROBLEM

Luis Ernesto Torres Guardia (UFF )
tepletg@vm.uff.br

Over the past several decades, stochastic programming models have gained wide acceptance by practitioners of optimization and operations research. Some research in financial engineering areas such as portfolio planning, in manufacturing networks and supply chain optimization, as well as in energy planning, as mentioned in Valente (2009), has demonstrated that the solutions of stochastic programming models are superiors of those of deterministic models in terms of expectation over future events.

Unfortunately, the size of stochastic programming problems grows exponentially with the number of decision stages considered in a model and with the number of possible outcomes of the random parameters at each stage. This is a challenging issue both in terms of modeling and solution algorithms. The implementation of large-scale mathematical programming models requires the generation of matrices that represent the problem constraints and that are stored in machine to be interpreted by solvers. This generation used to be performed by software modules and required considerable development effort.

Assuming that we have a finite number of scenarios, the classes of models, the multistage are essentially a huge size deterministic equivalent linear program (LP), and in principle, we can use any appropriate algorithms for solving this LP model, even the most efficient solvers do not scale up effectively with growth in stochastic programming problem size due to the large number of scenarios. With the rapid developments in designing and implementing interior point methods, in recent years have led to a major breakthrough in the field of the optimization and operation research. A characteristic property of interior point methods is that they typically require only a small number of iterative steps to reach a high precision solution, almost regardless of the size of the problem. This properly is certainty of critical importance in solving large scale stochastic programming problems. However, the key difficulty of the interior point methods is that, in every iteration step, usually involves solving a very large, perhaps even ill-conditioned, direction-finding linear system, named normal equation.

For the multi-stage stochastic network formulation of the dynamic
distribution problem and applying the interior point method, we use the Cholesky decomposition, implemented in MATLAB code.

Palavras-chaves: Linear programming, logistica, distribution stochastic
1. Introduction

Over the past several decades, stochastic programming models have gained wide acceptance by practitioners of optimization and operations research. Some research in financial engineering areas such as portfolio planning, in manufacturing networks and supply chain optimization, as well as in energy planning, as mentioned in Valente (2009), has demonstrated that the solutions of stochastic programming models are superiors of those of deterministic models in terms of expectation over future events.

Unfortunately, the size of stochastic programming problems grows exponentially with the number of decision stages considered in a model and with the number of possible outcomes of the random parameters at each stage. This is a challenging issue both in terms of modeling and solution algorithms. The implementation of large-scale mathematical programming models requires the generation of matrices that represent the problem constraints and that are stored in machine to be interpreted by solvers. This generation used to be performed by software modules and required considerable development effort.

Stochastic programming models are characterized by the presence of uncertainty associated with some of the model parameters. It is appropriate to consider these parameters to be random variables with specified probability distributions.

Optimum decision problems under uncertainty may be posed as two-stage or multistage scenario-based recourse problems. For complete, and others problems discussions, of these classes of models, readers are referred to the book of Birge and Louveaux (1997), Alonso-Ayuso, et al. (2009), Infanger (1994), Kall and Wallace (1994), Shapiro et al. (2009). Assuming that we have a finite number of scenarios, the classes of models, two-stage or multistage, are essentially a huge size deterministic equivalent linear program (LP), and in principle, we can use any appropriate algorithms for solving this LP model, even the most efficient solvers do not scale up effectively with growth in stochastic programming problem size due to the large number of scenarios. A successful solution method for this stochastic programming should exploit the special structure of the problem in order to cut down computational time. For this purpose, most of the solution methods in this area are based on specialized decomposition as mentioned in Berkelaar, et al. (2005).
With the rapid developments in designing and implementing interior point methods, in recent years have led to a major breakthrough in the field of the optimization and operation research, see for examples, Vanderbei (2008), Wright (1997). A characteristic property of interior point methods is that they typically require only a small number of iterative steps to reach a high precision solution, almost regardless of the size of the problem. This properly is certainty of critical importance in solving large scale stochastic programming problems. However, the key difficulty of the interior point methods is that, in every iteration step, usually involves solving a very large, perhaps even ill-conditioned, direction-finding linear system, named normal equation. Some others variants of the interior point methods of finding search-direction can be used for solving multistage stochastic linear programs. For example, Liu and Sun (2004) proposed of version of the primal-dual potential reduction algorithm for solving that problem.

In this work, we propose a new decomposition method for the multistage stochastic distribution problem with uncertain demands which involve the allocation of goods or resources to storage areas in anticipation of forecasted markets demands. This class of model is discussed in Cheung and Powell (1996), Valente et al. (2009) and they presented some examples that fall within the framework of distribution planning. They proposed a two-stage and multi-stage formulations of this distribution problems using the framework of dynamic networks with random arc capacities. For other side, for finding the solution of a multistage stochastic programming problem, Sun and Liu (2000) use the homogeneous self-dual interior point method which is particularly suitable for computing the Newton directions, which are determined using a 3-step decomposition approach. Liu and Fukushima (2006) propose a parallelizable preprocessing method, which exploits effectively the structure for some production planning problem. They solve some test problems running a MATLAB function, which implies that a primal-dual interior-point method is used to solve those problems.

For the multi-stage stochastic network formulation of the dynamic distribution problem and applying the interior point method, we use the Cholesky decomposition see for example, Mészáros (2005) for solving the corresponding normal equations. Others types of decomposition can be used, like the AINV method, see Bellavia, et al. (2013) for solving the crucial normal equations. As an alternative to the normal equation approach, some authors use a set of equations usually referred to as the augmented system and in some cases of the structure of the set of the
constraints of the LP, the augmented system improves on normal equations. We will study this approach in a future research.

This paper is organized as follows. Section 2, we present the basic ideas of the interior point technique to be employed in our approach. In section 3 we shall introduce the multi-stage stochastic linear programming problem for the distribution problem with uncertain demands. In section 4, we discuss how to solve the direction-finding linear system. In section 5 we shall present some numerical results for using the decomposition method. Finally, we conclude the paper in section 6.

2. Interior point method

The common experience of extensive numerical tests indicates that the primal – dual methods from the family of interior point methods are the most powerful ones in practice. In this section, we do not describe any of their variants in details. Readers interested in the full development of these algorithms are referred to the books of Wright(1997), Vanderbei (2008).

We consider the following primal linear programming problem:

$$\min c^T x$$

subject to

$$Ax = b,$$

$$x \geq 0,$$

where A is an m x n matrix, c and x are vectors, and b is an m vector. We assume that A is of full row rank.

Let y denote the m vector of dual variables. The dual problem associated with the primal problem can be written as follows:

$$\max b^T y$$

subject to

$$A^T y + z = c,$$

$$z \geq 0,$$

where z is the n vector of dual slack variables. Let X denote the diagonal matrix whose diagonal elements are the components of x, that is, X = diag(x) and, in a similar manner, let Z = diag(z).

The primal –dual algorithm generates a sequence of primal - dual solutions, which converges to the optimum. The Newton method is used in the primal –dual algorithm, and in this case, the following linear system, named normal equations, is obtained:

$$(AZ^{-1}XA^T) dy = AZ^{-1}X(\zeta_c - X^{-1} \zeta_\mu) + \zeta_b.$$
\[
\begin{align*}
\text{dz} &= \zeta_c - A^T \text{dy}, \\
\text{dx} &= Z^{-1}(\zeta_{\mu} - \text{Xdz}),
\end{align*}
\]

where \( \text{dw} = (\text{dx}, \text{dy}, \text{dz})^T \) is the Newton’s direction and \( \zeta_b = b - Ax, \zeta_c = c - A^T y - z, \zeta_{\mu} = \mu e - Xze, e = (1, \ldots, 1) \in \mathbb{R}^n \).

Given an iterate \( w = (x,y,z) \), we choose a step-length \( \alpha > 0 \) such that \( w = w + \beta \alpha \text{dw} \), \( \alpha \) is determined by a suitable line search procedure and \( \beta \in (0, 1) \) and near 1. Repeat the procedure until a given precision is reached.

Since \( (AZ^{-1}XA^T) \) is a positive definite matrix, the well-known Cholesky factorization can be used to solve the normal equations, but difficulties can arise when \( A \) contains dense columns. The matrix \( (AZ^{-1}XA^T) \) can be factorized in the following way: \( (AZ^{-1}XA^T) = RR^T \), where \( R \) is a lower triangular matrix. Solving the normal equations is equivalent to solve two triangular systems of linear equations, for more details of the factorization method, see the book of Nocedal and Wright (1999).

### 3. Multi-stage distribution problem

In this work, we study the dynamic distribution problems which require decisions to be made over the horizontal time. For simplicity, we assume that uncertainty only occurs in the demand for the product in each of the future time periods. Consider then a distribution planning horizon where a stage is defined as a period of time during which the customer demands are realized. We represent by \( N \) the number of stages.

Before presenting a multi-stage distribution problem with uncertain demands, we first introduce some notation that will be used throughout this work.

Let \( \xi \) be a random vector defined over a probability space \( (\Omega, \mathcal{F}, \mathcal{P}) \) where \( \Omega \) is the set of elementary outcomes \( w \), \( \mathcal{F} \) is the event space and \( \mathcal{P} \) is the probability measure. In this model, an elementary outcome \( w \in \Omega \) consists of a set of outcomes \( w_1, \ldots, w_N \) where \( w_i \) represents an outcome in stage \( t \). We denote by \( \xi(t) \) the vector of customer demands in stage \( t \). We have then the following notation:

Indices and Parameters:

\( P: \) set of indices representing plants, with \( i \in P; \)
W: set of indices representing warehouses, with \( j \in W \);

C: set of indices representing customers, with \( k \in C \);

Nc: set of indices representing scenarios, with \( lc \in Nc \);

\( c_{ij}(t) \): cost of shipping a unit of product from plant \( i \) to warehouse \( j \) in stage \( t \);

\( c_{ii}(t) \): holding cost per unit of product in plant \( i \) in stage \( t \);

\( q_{jk}(t) \): cost of shipping a unit of product from warehouse \( j \) to customer \( k \) in stage \( t \);

\( q_{jj}(t) \): holding cost per unit of product in warehouse \( j \) in stage \( t \);

\( R_i(t) \): amount of goods product at plant \( i \) in stage \( t \).

Decision Variables

\( x_{ij}(t) \): amount of goods shipped from plant \( i \) to warehouse \( j \) in stage \( t \);

\( x_{ii}(t) \): inventory in plant \( i \) in stage \( t \);

\( y_{jk}(t,w_i) \): amount of products being shipped from warehouse \( j \) to customer \( k \) in stage \( t \) when the realization of customer demands is \( \xi(t,w_i) \);

\( y_{jj}(t) \): inventory in warehouse \( j \) in stage \( t \).

We assume there is no inventory in the plants at the beginning of the planning horizon and all the demands are satisfied.

We set out below the mathematical formulation of the multi-period stochastic linear programming problem for the distribution model, based on the work of Cheung and Powell (1996) and Valente et al. (2008):

\[
\begin{align*}
\min & \quad \sum_{i \in I} \sum_{j \in J} \sum_{c \in N_c} c_{ij}(t)x_{ij}(t) + \sum_{j \in J} \sum_{k \in K} \sum_{n \in N_n} q_{jk}(t)y_{jk}(t,w_i) \\
\text{subject to:} & \quad \sum_{i \in I} x_{ij}(1) + x_{ii}(1) = R_i(1), \quad \text{(2)} \\
& \quad \sum_{i \in I} x_{ij}(t) + x_{ii}(t) - x_{ii}(t-1) = R_i(t), \quad t \neq 1, \quad \text{(3)} \\
& \quad - \sum_{i \in I} x_{ij}(1) + y_{ij}(1) + y_{jk}(1,w_i) = 0, \quad \text{(4)} \\
& \quad - \sum_{i \in I} x_{ij}(t) - y_{ij}(t-1) + y_{ij}(t) + y_{jk}(t,w_i) = 0, \quad t \neq 1, \quad \text{(5)}
\end{align*}
\]
\[ \sum_{i,j} x_{ij} - x_{ii}(N-1) = R_i(N), \]  \hspace{1cm} (6)

\[ - \sum_{i,j} x_{ij} - y_{jj}(N-1) + y_{jk}(N,w_N) = 0, \]  \hspace{1cm} (7)

\[ y_{j,k}(t,w_i) \leq \xi_k(t,w_i), \]  \hspace{1cm} (8)

\[ x_{ij}(t) \geq 0, \quad y_{jk}(t,w_i) \geq 0. \]  \hspace{1cm} (9)

To illustrate the dynamic distribution model with stochastic demands with an \( N \) – stage planning horizon, consider the problem depicted in the following Figure 1.

**Figure 1. Multistage stochastic distribution problem**

4. Implementation

In this section, we present the implementation of the interior point method to determine the Newton’s direction, using the normal equations. For this, we note that the sets of constraints (2) – (7) represent the usual flow conservation conditions. By adding the corresponding slack variable to the constraint (8), it is now an equality constraint and the stochastic distribution model is written as the following linear programming problem:

\[ \min (\bar{c})^T \bar{x} \]  \hspace{1cm} (10)

subject to: \[ \bar{A} \bar{x} = \bar{b} \]  \hspace{1cm} (11)
where:

\[
\overline{A} = \begin{bmatrix}
A_1 & B_1 & A_i \\
B_1 & A_i & B_1 \\
 & & A_i \\
 & & B_1 \\
\end{bmatrix}, \quad \overline{A} \in \mathbb{R}^{[N\times m_1] \times [(N-1) \times m_2 + m_3]}
\]

where we denote \( m_1 = m + n \times n_c + p \times p_c \); \( m_2 = m + n + m \times n + n \times p \times n_c + p \times n_c \); \( m_3 = m_2 - m - n \),

The matrix \( A_1 \in \mathbb{R}^{m_1 \times m_2} \) is given in the following form:

\[
A_1 = \begin{bmatrix}
I_m & A \\
I_n & C & B \\
\end{bmatrix}
\]

\begin{bmatrix}
I_n & C & B \\
E & I_p \\
E & I_p \\
\end{bmatrix}

I_m, I_n and I_p denote the \( m \times m \), \( n \times n \) and \( p \times p \) identity matrices respectively.

Now, matrix \( A \in \mathbb{R}^{m \times (m \times n)} \) has the following form:

\[
A = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & & \\
& 1 & 1 & \\
& & 1 & 1 \\
\end{bmatrix}
\]

Matrix \( B \in \mathbb{R}^{n \times (n \times p)} \) has the same form of matrix \( A \).

Matrix \( C \in \mathbb{R}^{n \times (n \times n)} \) has the form:

\[
C = \begin{bmatrix}
-1 & -1 \\
& -1 \\
& & -1 \\
& & & -1 \\
\end{bmatrix}
\]

Matrix \( E \in \mathbb{R}^{p \times (n \times p)} \) has the same form of matrix \( C \) such that \( E = -C \).
Now, the matrix $B_1$ is given by the form:

$$
B_1 = \begin{bmatrix}
-I_m & 0 & 0 \\
0 & -I_n & 0 \\
0 & 0 & -I_n \\
0 & 0 & 0
\end{bmatrix}, \quad B_1 \in \mathbb{R}^{m_1 \times m_2}
$$

Matrix $C_1$ has the form:

$$
C_1 = \begin{bmatrix}
A & C & B \\
C & B & I_p \\
E & I_p
\end{bmatrix}, \quad C_1 \in \mathbb{R}^{m_1 \times m_3}
$$

Vector $\bar{x}$ represents an overall vector, $\bar{x} \in \mathbb{R}^{(N-1) \times m_2 + m_3}$ where $\bar{x} = (x^1, \ldots, x^N)$ and $x^i$ is its sub-vector, $x^i \in \mathbb{R}^{m_2}$, $t=1, \ldots, N-1$, $x^N \in \mathbb{R}^{m_3}$,

$x^1 = (x_{i1}(t), y_{i1}(t), x(t), y(t), y(t)); x^N = (x(N), y(N), y(N))$

$x_i(t) = (x_{1i}(t), \ldots, x_{mi}(t)), t=1, \ldots, N-1$;

$y_i(t) = (y_{1i}(t), \ldots, y_{mi}(t)), t=1, \ldots, N-1$;

$x(t) = (x_{ij}(t)), i=1, \ldots, m; j=1, \ldots, n; t=1, 1, \ldots, N$;

$y(t) = (y_{jkl}(t)), j=1, \ldots, n; k=1, \ldots, p; l=1, 1, \ldots, N; t=1, \ldots, N$;

$y(t) = (1, \ldots, 1) \in \mathbb{R}^p$.

Vector $\bar{b} = (b_1, \ldots, b_N)^T \in \mathbb{R}^{N \times m_1}$ where $b_i = (R_i(t), \ldots, R_m(t), \xi(t,w_1), \ldots, \xi(t,w_{nc})), b_i \in \mathbb{R}^{m_1}$.

Applying the interior point method to the multi-stage stochastic distribution model defined as the linear programming problem (10) – (12), now as a deterministic model, using scenario approach, we need to solve the normal equations and for this, it is performed block multiplications to determine the matrix $(\bar{A}Z^{-1}X\bar{A}^T)$, which has the following structure:
(A Z⁻¹ X Aᵀ) =

\[
\begin{bmatrix}
A_{11} & A_{12} & & & \\
A_{12}^T & A_{22} & A_{23} & & \\
& A_{23}^T & A_{33} & & \\
& & & \cdots & \\
& & & & A_{N,N-1}
\end{bmatrix}
\]

where

\[
A_{11} = A₁D₁A₁^T,
A_{ii} = B₁Dᵢ⁻¹Bᵢ⁺¹ + A₁DᵢAᵢ^T,
A_{NN} = B₁D_{N⁻¹}B₁⁺¹ + CᵢDᵢCᵢ^T,
A_{ij} = A₁Dⱼ⁻¹Bⱼ⁺¹, j = 2, ..., N
\]

and \(Dᵢ\) a diagonal matrix given by \(Dᵢ = Zᵢ⁻¹Xᵢ\), \(i = 1, ..., N\).

Now, the solution of the corresponding normal equations:

\((A Z⁻¹ X Aᵀ) dy = r,\)

where \(r = AZ⁻¹X(ζₖ - X⁻¹ζₖ) + ζₖ\),

is, after some matrix operations, given by:

\(A_{11} dy = r₁ - A_{12} dy₂,\)
\((A_{ii} - A_{ii}ᵀ Bᵢ⁻¹,₁⁻¹) dy = rᵢ - A_{ii}ᵀ bᵢ⁻¹ - A₁,ᵢ⁺¹ dyᵢ⁺¹,\)
\((A_{NN} - A₁,₁⁻¹ᵀ Bₖ⁻¹,₁⁻¹) dy = rₖ - A₁,₁⁻¹ᵀ bₖ⁻¹,\)

where \(dy = (dy₁, dy₂, ..., dy_N)ᵀ\) and \(r = (r₁, r₂, ..., rₖ)ᵀ\),

and the following matrices and vectors can be found in the following way:

\(B₁₁ = A₁⁻¹ᵀ B₁₂ᵀ,\)
\(Bⱼⱼ = (Aⱼⱼ⁻¹ᵀ Bⱼ⁻¹,₁⁻¹)⁻¹ A₁,ⱼ⁺¹, j = 2, ..., N⁻¹,\)
\(b₁ = A₁⁻¹ᵀ r₁,\)
\(bⱼ = (Aⱼⱼ⁻¹ᵀ Bⱼ⁻¹,₁⁻¹)⁻¹ (rⱼ - A₁ᵀ bⱼ⁻¹), j = 2, ..., N⁻¹.\)

As it can be seen that to find the vector \(dyᵢ\), \(i = 1, ..., N\), it is necessary to determine the inverse of some matrices, see Bellavia et al. (2013), or use the Cholesky decomposition. In this later case, for example, matrix \((A_{ii} - A_{ii}ᵀ B₁⁻¹,₁⁻¹)\) can be decomposed in the following manner:

\((A_{ii} - A_{ii}ᵀ B₁⁻¹,₁⁻¹) = Rᵢ Rᵢᵀ,\)

where, \(Rᵢ\) is a lower triangular matrix.

With the above decomposition, now it is possible to determine the matrices \(Bᵢᵢ\), \(i = 1, ..., N⁻¹,\) and
the vectors $b_i$, $i = 1, \ldots, N-1$, and in this case it is not necessary to find the corresponding inverse matrices associated with those matrices and vectors. Instead of using the Cholesky decomposition, it could be used the AINV method, see Benzi et al. (2000), and in this case, the inverse of the above matrices is approximately determined directly, that is for example:

$$(A_{ii} - A_{1i}^T B_{i-1,i-1}^{-1})^{-1} \approx Z_i P_i^{-1} Z_i^T,$$

where $Z_i$ is an upper triangular matrix with a diagonal of 1’s and $P_i$ is a diagonal matrix. This procedure will be studied in the future work.

4. Computational Results

In order to see how the primal-dual interior point method proposed in this work, in this section we present the numerical results to solve the multi-stage stochastic distribution model for several instances of different sizes. The vector $\xi(t)$ of customer demands in stage $t$ is given in such a way that all the demands are satisfied. For simplicity, the cost vector associated to the function objective of the primal problem, that is $\overline{c}$, is a vector of 1’s.

The code is written in MATLAB and run on a microcomputer core i5 with 2.53 GHz processor and 4 GB of memory under Windows 7 operating system. A starting point is given in the primal-dual method and it may not be viable. It was used in all computational tests the value $\beta = 0.99995$.

The corresponding system of normal equations that appears in the context of interior point methods is solved using the Cholesky decomposition. This subroutine is available in the MATLAB code. The algorithm terminates when the following relative gap, computed by the formula:

$$\frac{|(\overline{c})^T \overline{x} - (\overline{b})^T \overline{y}|}{1.0 + |(\overline{b})^T \overline{y}|}$$

is less than $10^{-4}$, where $(\overline{c})^T \overline{x}$ is the value of the objective function of the primal linear problem and $(\overline{b})^T \overline{y}$ 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 1, named (iter), now, $\overline{c}$ denoted the value of the objective function of the primal linear
programming problem, that is, \((\tilde{c})^T \tilde{x}\) and \(f_d\) is the value of the objective function of the corresponding dual problem, that is, \((\tilde{b})^T \tilde{y}\). The time was measured with the cpu-time function of MATLAB and is recorded in the last line of the Table 1.

In the following Table 1, we summarize preliminary numerical results for some test problems.

<table>
<thead>
<tr>
<th>m</th>
<th>2</th>
<th>2</th>
<th>4</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>p</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>nc</td>
<td>15</td>
<td>21</td>
<td>12</td>
<td>60</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>(f_d)</td>
<td>(1.886e^3)</td>
<td>(2.485e^3)</td>
<td>(2.220e^3)</td>
<td>(8.700e^3)</td>
</tr>
<tr>
<td>(\mu)</td>
<td>(6.2485e^{-3})</td>
<td>(3.9126e^{-3})</td>
<td>(1.3770e^{-6})</td>
<td>(5.6103e^{-6})</td>
</tr>
<tr>
<td>(\text{iter})</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>(\text{time})</td>
<td>0.049689</td>
<td>0.077512</td>
<td>0.146380</td>
<td>3.225833</td>
</tr>
</tbody>
</table>

6. Conclusions

In this paper, it is proposed the primal-dual interior-point method to solve the multistage stochastic distribution model formulated as a deterministic linear programming problem via scenarios, specifically for the distribution problem with stochastic demand. Generally, the stochastic problem grows rapidly as the number of scenarios grows. For this reason, it is used the interior point method to solve the new large scale linear program. The solution of the associated linear system of the interior-point method is determined using the traditional Cholesky decomposition. For this multistage stochastic model, all the matrices associated to the linear program are stored without using any technique of exploiting the structure of these matrices. This will be a subject of future research. The numerical results report the efficiency of the primal-dual interior-point method for solving the multistage stochastic distribution model.

References


