An Algorithm for Computing Minimal Coxian Representations

Qi-Ming He
Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia B3J 2X4, Canada, qi-ming.he@dal.ca

Hanqin Zhang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100080 Beijing, People’s Republic of China, hanqin@amt.ac.cn

This paper presents an algorithm for computing minimal ordered Coxian representations of phase-type distributions whose Laplace-Stieltjes transform has only real poles. We first identify a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal with respect to the number of phases involved. The conditions establish a relationship between the Coxian representations of a Coxian distribution and the derivatives of its distribution function at zero. Based on the conditions, the algorithm is developed. Three numerical examples show the effectiveness of the algorithm and some geometric properties associated with ordered Coxian representations.

Key words: Coxian distribution; phase-type distribution; matrix-exponential distribution; matrix-analytic methods; nonlinear programming

History: Accepted by Edward P. C. Kao, Area Editor for Computational Probability and Analysis; received August 2005; revised May 2006; accepted April 2007. Published online in Articles in Advance November 26, 2007.

1. Introduction

Coxian distributions have found many applications in the study of queueing, reliability, supply chains, insurance and risk, and telecommunications (Asmussen 2000, 2003; Cox 1955a, b; Feldmann and Whitt 1998; Haddad et al. 1998; Latouche and Ramaswami 1999; Neuts 1981, 1989; Sasaki et al. 2004). For instance, Coxian distributions have been used to model service times and interarrival times in queueing models, component life times in reliability models, and interarrival times of claims in insurance and risk models. Analysis of these models usually involves complicated computational procedures using detailed information about Coxian distributions. In particular, the numbers of phases of these distributions play a significant role in computation. Consequently, reduction in the number of phases of Coxian representations can improve efficiency in computation and performance analysis. A number of studies have been carried out on Coxian and related distributions (e.g., Cox 1955a, b; Cumani 1982; Dehon and Latouche 1982; Harris et al. 1992; Heijden 1988; Mocanu and Commault 1999; O’Cinneide 1989, 1991, 1993; Osogami and Harchol-Balter 2003b).

Neuts (1975) generalized Coxian distributions into phase-type (PH) distributions as the distribution of the absorption time of a finite state Markov process, which made it possible to study complicated queueing models such as the PH/PH/c queue analytically and numerically (Takahashi 1981). It is well known that the representation of a PH-distribution is not unique. To reduce the time complexity of algorithms associated with PH-distributions, it is useful to find a PH-representation with the minimal number of phases for a PH-distribution, known as the minimal PH-representation problem (Aldous and Shepp 1987; Commault and Mocanu 2003; Mocanu and Commault 1999; O’Cinneide 1989, 1990, 1991, 1993; Osogami and Harchol-Balter 2003a). Chapter 2 in Neuts (1981) provided historical notes on PH-distributions. Commault and Mocanu (2003) and O’Cinneide (1999) reviewed studies on PH-distributions.

While the problem of finding a minimal PH-representation for a PH-distribution is still open, the problem of finding simpler representations for PH-distributions or for some subsets of PH-distributions has been studied extensively (Bobbio et al. 2002, 2004; Commault and Mocanu 2003; Cumani 1982; Dehon and Latouche 1982; He and Zhang 2006a, b; Mocanu and Commault 1999; O’Cinneide 1990, 1991, 1999). Coxian representation is one of the simpler PH-representations being investigated. Cumani (1982) showed that any PH-distribution with a triangular PH-representation is Coxian and a Coxian representation of the same order can be found. O’Cinneide (1989, 1991, 1993) introduced a number of new
concepts in the study of PH and Coxian distributions. Concepts such as PH-simplicity and triangular order shall be used in this paper. O’Cinneide (1991) proved that all PH-distributions whose Laplace-Stieltjes transforms have only real poles are Coxian. In O’Cinneide (1993), a necessary and sufficient condition was given for a Coxian representation to be minimal.

Approximating general probability distributions with Coxian or PH-distributions has been investigated extensively as well. Johnson (1993) and Johnson and Taaffe (1989; 1990a, b) introduced several methods for fitting PH and Coxian distributions to general distributions by matching their lower moments. Osogami and Harchol-Balter (2003b) focused on Coxian approximations of distributions. Algorithms were introduced for computing Coxian distributions matching the first three moments of the original probability distributions. Their algorithms may find Coxian representations with the minimal number of phases, though it is not guaranteed. Sasaki et al. (2004) developed an approximation method for finding Coxian distributions as approximations of general probability distributions. Other works on Coxian approximations of distributions are Asmussen et al. (1996), Feldmann and Whitt (1998), and Heijden (1988).

Coxian distributions were studied as a subset of PH-distributions and as approximations to general distributions. However, no algorithm for computing a minimal Coxian representation has been developed. We find a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal. This set of conditions establishes an explicit relationship between the parameters of a Coxian representation of a Coxian distribution and the derivatives of its distribution function at zero. Based on these conditions, an algorithm is introduced for computing a minimal Coxian representation for any Coxian distribution or, equivalently, any PH-distribution whose Laplace-Stieltjes transform has only real poles. Our necessary and sufficient conditions are based on the spectral polynomial method, which was developed in He and Zhang (2006b) (also see He and Zhang 2005, 2006a) for computing bidiagonal representations of PH-distributions and matrix-exponential distributions. Use of spectral polynomials and the derivatives of the distribution function at zero distinguishes our method.

Section 2 gives definitions and some preliminary results for development of the algorithm. Section 3 shows how to compute the algebraic degree of a Coxian distribution, which is a lower bound on the triangular order of Coxian distribution. In §4, we present a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal.

In §5, we show that a minimal ordered Coxian representation can be found by solving a series of nonlinear programs for any PH-distribution whose Laplace-Stieltjes transform has only real poles. Three numerical examples are presented in §6 to show the effectiveness of the algorithm and some geometric properties associated with Coxian distributions. Finally, in §7, we discuss future research.

2. PH, ME, and Coxian Distributions
A square matrix $T$ with negative diagonal elements, nonnegative off-diagonal elements, nonpositive row sums, and at least one negative row sum is called a sub-generator in the general literature of Markov process. We shall call a sub-generator $T$ of finite size a PH-generator. Define an infinitesimal generator for a continuous-time Markov chain with $m + 1$ states

$$
\begin{pmatrix}
T & -Te \\
0 & 0
\end{pmatrix},
$$

where state $m + 1$ is an absorption state and $e$ is the column vector with all elements being one. The matrix $T$ is an $m \times m$ PH-generator. We assume that states $\{1, 2, \ldots, m\}$ are transient. Let $\alpha$ be a nonnegative vector of size $m$ for which the sum of its elements is at least one. We call the distribution of the absorption time of the Markov chain to state $m + 1$, with initial distribution $(\alpha, 1 - \alpha e)$, a phase-type distribution (PH-distribution). We call the pair $(\alpha, T)$ a PH-representation of that PH-distribution. The number $m$ is the order of the PH-representation $(\alpha, T)$. The probability distribution function of the PH-distribution is given as $1 - \alpha \exp[Tt]e$ for $t \geq 0$, and the density function is given as $-\alpha \exp[Tt]Te$ for $t \geq 0$. If $\alpha e = 0$, the distribution has a unit mass at time 0. There is no need for a PH-representation for such a distribution. If $\alpha e \neq 0$, the expression $\alpha \exp[Tt]e$ can be written as $\alpha e (\alpha/(\alpha e)) \exp[Tt]e$.

Thus, the study on the representation of $(\alpha, T)$ is equivalent to that of $(\alpha/(\alpha e), T)$. Without loss of generality, we assume that $\alpha$ is a vector for which the sum of all its elements is one. That implies that all probability distributions we consider have a zero mass at $t = 0$. See Chapter 2 in Neuts (1981) for basic properties of PH-distributions.

It is possible that $1 - \alpha \exp[Tt]u$ is a probability distribution function for a row vector $\alpha$ of size $m$, an $m \times m$ matrix $T$, and a column vector $u$ of size $m$, where the elements of $\alpha$, $T$, and $u$ can be complex numbers. For this case, the 3-tuple $(\alpha, T, u)$ is called a matrix-exponential representation (ME-representation) of a matrix-exponential distribution (ME-distribution). Without loss of generality, we assume that $\alpha u = 1$. Apparently, the class of PH-representations is a subset of the class of ME-representations. See Asmussen and
Bladt (1996) and Lipsky (1992) for more details about ME-distributions and their applications in queueing theory.

Throughout this paper, when \((\alpha, T)\) is used, \(\alpha\) is nonnegative, \(T\) is a PH-generator, and \((\alpha, T)\) is a PH-representation of a PH-distribution. When \((\alpha, T, u)\) is used, \(\alpha\) may not be nonnegative and \(T\) may not be a PH-generator.

For \(x = (x_1, x_2, \ldots, x_N)\) where \(N\) is a positive integer, a bidiagonal matrix \(S(x)\) is defined as

\[
S(x) = \begin{pmatrix}
-x_1 & 0 & \cdots & 0 \\
x_2 & -x_2 & \cdots & \vdots \\
0 & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & x_{N-1} & -x_{N-1} & 0 \\
0 & \cdots & 0 & x_N & -x_N
\end{pmatrix}
\]

If \(S(x)\) is a PH-generator, it is called a bidiagonal PH-generator. If \([x_1, x_2, \ldots, x_N]\) are all real and positive and \(\beta\) is a probability vector (i.e., \(\beta \geq 0\) and \(\beta e = 1\)), then \((\beta, S(x))\) is called a Coxian representation, which represents a Coxian distribution. The class of Coxian distributions is a subset of the class of PH-distributions. Further, if \(x_1 \geq x_2 \geq \cdots \geq x_N > 0\), then \((\beta, S(x))\) is called an ordered Coxian representation. The following theorem characterizes the class of Coxian distributions.

**Theorem 1** (Theorem 4.1 in O’Cinneide 1991 and Theorem 5.2 in O’Cinneide 1993). Every PH-distribution whose Laplace-Stieltjes transform has only real poles is a Coxian distribution and has an ordered Coxian representation. □

Theorem 1 implies that the class of (standard) Coxian distributions defined by (1) covers many well-known probability distributions. For instance, generalized Erlang distributions, mixtures of exponential distributions, PH-distributions with a triangular PH-generator (Cumani 1982), and PH-distributions with a symmetric PH-generator (He and Zhang 2006b) are all special cases of Coxian distributions.

It is well known that the PH-representation of a PH-distribution and the ordered Coxian representation of a Coxian distribution are not unique. A PH-representation with the minimal number of phases is called a minimal PH-representation. The number of phases of a minimal PH-representation is called the PH-order of the corresponding PH-distribution. An ordered Coxian representation with the minimal number of phases is called a minimal ordered Coxian representation. The number of phases of a minimal ordered Coxian representation is called the triangular order of the corresponding PH-distribution. Theorem 6.2 in O’Cinneide (1993) gave a necessary and sufficient condition for an ordered Coxian representation to be minimal. In the following, a new set of necessary and sufficient conditions is obtained for an ordered Coxian representation to be minimal, and a series of nonlinear programs is developed for computing a minimal ordered Coxian representation.

In general, a PH-representation that represents a Coxian distribution has Coxian representations, but the orders of the Coxian representations may be larger than the order of the PH-representation (see Theorem 4.5 in He and Zhang 2006b and Example 1 in this paper).

The following spectral polynomial algorithm introduced in He and Zhang 2006b is useful for computing bidiagonal representations of ME-distributions and PH-distributions. Let \((\alpha, T, u)\) be an ME-representation of order \(m\). Denote by \(x = (x_1, x_2, \ldots, x_N)\) a vector whose elements are complex numbers, where \(N\) is a positive integer. If all elements of \(x\) are nonzero, define

\[
\begin{align*}
|p_1 &= -Tu/x_1; \\
|p_n &= (x_{n-1}I + T)p_{n-1}/x_n, \quad 2 \leq n \leq N; \\
|p_{N+1} &= (x_NI + T)p_N.
\end{align*}
\]

Let \(P = (p_1, p_2, \ldots, p_N)\), which is an \(m \times N\) matrix. Similarly to the proof of Proposition 3.1 in He and Zhang (2006b), the following results can be proved.

**Theorem 2.** If \(p_{N+1} = 0\), we have \(TP = PS(x)\) and \(Pe = u\). Further, the matrix-exponential representation \((\alpha, T, u)\) has an equivalent bidiagonal representation \((\beta, S(x), e)\) of order \(N\) with \(\beta e = \alpha u = 1\) and \(\beta = \alpha P\). □

The above spectral polynomial algorithm is called the Post-T spectral polynomial algorithm in He and Zhang (2006b). We remove “Post-T” because it is nonessential.

For any two vectors \(x\) and \(y\), we say that \(x\) is part of \(y\) if any element that appears \(k (\geq 0)\) times in \(x\) appears at least \(k\) times in \(y\). A particular choice of \(x\) was given in He and Zhang (2006b). Denote by \([-\lambda_1, -\lambda_2, \ldots, -\lambda_m]\) the spectrum of \(T\) (counting multiplicities, i.e., with repeated elements). If \((\lambda_1, \lambda_2, \ldots, \lambda_m)\) is part of \(x\), by the Cayley-Hamilton theorem (Lancaster and Tismenetsky 1985), we must have \(p_{N+1} = 0\). If \(x_1 \geq x_2 \geq \cdots \geq x_N > 0\) and \(\alpha P\) is nonnegative, \((\alpha P, S(x))\) is an ordered Coxian representation. For this case, an ordered Coxian representation is obtained for \((\alpha, T, u)\). For more about the spectral polynomial algorithm, see He and Zhang (2005; 2006a, b).

In O’Cinneide (1989), a PH-generator \(T\) is called PH-simple if each PH-distribution \((\alpha, T)\) has a unique representation of the form \((\alpha, T)\), i.e., if \(\alpha_1\) and \(\alpha_2\) are two probability vectors and \(\alpha_1 \neq \alpha_2\), then \((\alpha_1, T)\) and \((\alpha_2, T)\) represent two different PH-distributions. By
Theorem 1 in O’Cinneide (1989), a PH-generator T is PH-simple if and only if \{TE, T^2e, \ldots, T^me\} are independent vectors. By Theorem 2 in O’Cinneide (1989), a PH-generator T is PH-simple if and only if it has no left eigenvector orthogonal to e.

The above definition of PH-simplicity can be generalized to ME-representations of the form (α, T, e), where T is a PH-generator, and αe = 1 (but α may not be nonnegative). That is, a PH-generator T is PH-simple if each ME-distribution (α, T, e) has a unique representation of the form (α, T, e). This generalization of simplicity is exactly the same as the property of simplicity for the case of a PH-generator. It is easy to check that both Theorems 1 and 2 in O’Cinneide (1989) are still true under the generalized definition of PH-simplicity. Next, we establish a relationship between PH-simplicity and the spectral polynomial algorithm.

**Theorem 3.** Assume that (1) T is PH-simple; (2) u = e; (3) all elements of x are nonzero; and (4) N = m. Then the matrix P obtained from (2) is invertible.

**Proof.** From (2), we obtain

\[ p_0 = c_{n, 1}Te + c_{n, 2}T^2e + \cdots + c_{n, n}T^ne, \quad 1 \leq n \leq m, \]

where \{c_{n, j}, 1 \leq j \leq n \leq m\} are some constants and \(c_{n, n} \neq 0\), for 1 \leq n \leq m. Since T is PH-simple, by Theorem 1 in O’Cinneide (1989), vectors Te, T^2e, \ldots, T^me are independent. Together with \(c_{n, n} \neq 0\), for 1 \leq n \leq m, it is easy to see that the vectors \(p_1, p_2, \ldots, p_m\) are independent, and consequently, the matrix P is invertible. \(\square\)

### 3. A Minimal ME-Representation

We assume that the Laplace-Stieltjes transform of a PH-distribution (α, T) of order m has only real poles. According to Theorem 1, (α, T) represents a Coxian distribution. We will develop a method for finding the triangular order of (α, T) and for computing a minimal ordered Coxian representation of (α, T). We begin with computation of a minimal ME-representation.

We assume that the PH-generator T has in total K distinct eigenvalues \{-\mu_1, -\mu_2, \ldots, -\mu_K\}. Assume that the algebraic multiplicity of the eigenvalue \(-\mu_k\) is \(m_k\), i.e., \(-\mu_k\) appears a total of \(m_k\) times in the vector \((-\alpha_1, -\alpha_2, \ldots, -\alpha_m)\). Let \(\mu = (\mu_1, \mu_2, \ldots, \mu_K)\) and \(m = (m_1, m_2, \ldots, m_K)\). Define

\[ S(\mu, m) = \begin{pmatrix}
E(\mu_1, m_1) & & \\
& \ddots & \\
& & E(\mu_K, m_K)
\end{pmatrix}, \]

where the matrix \(E(\mu_k, m_k) = S((\mu_k, \mu_k, \ldots, \mu_k))\) is of size \(m_k, 1 \leq k \leq K\). Note that if \(\mu_k\) is positive real,

\[ E(\mu_k, m_k) \]

is the PH-generator of an Erlang distribution of order \(m_k\) with parameter \(\mu_k\).

**Lemma 1.** If \(\mu_1, \mu_2, \ldots, \mu_K\) are distinct, then no left eigenvector of \(S(\mu, m)\) corresponding to eigenvalue \(-\mu_k\) is orthogonal to the vector e. Furthermore, if \(\mu_1, \mu_2, \ldots, \mu_K\) are distinct positive real numbers, then the PH-generator \(S(\mu, m)\) is PH-simple.

**Proof.** Suppose that some \(v\) is a left eigenvector of \(S(\mu, m)\) corresponding to eigenvalue \(-\mu_k\). Since \(\mu_1, \mu_2, \ldots, \mu_K\) are distinct, \(v\) has structure (0, 0, 0, \(v_k, 0, \ldots, 0\)), where \(v_k\) is an eigenvector of the matrix \(E(\mu_k, m_k)\). Post-multiplying \(e\) on both sides of \(-\mu_k v_k = v_k E(\mu_k, m_k)\), we obtain

\[ -\mu_k v_k e = v_k E(\mu_k, m_k) e = -\mu_k v_k (1, 0, \ldots, 0)', \]

where \((1, 0, \ldots, 0)’\) is the transpose of \((1, 0, \ldots, 0)\). If \(v e = 0\), Equation (3) implies that the first element of \(v_k\) is zero. From \(-\mu_k v_k = v_k E(\mu_k, m_k)\) it can be shown that the vector \(v_k\) is zero. Thus, \(v\) is zero, which is a contradiction. Therefore, no left eigenvector of \(S(\mu, m)\) is orthogonal to e.

If \(\mu_1, \mu_2, \ldots, \mu_K\) are distinct positive real numbers, then \(S(\mu, m)\) is a PH-generator. Because no left eigenvector of \(S(\mu, m)\) is orthogonal to e, by Theorem 2 in O’Cinneide (1989), \(S(\mu, m)\) is PH-simple. \(\square\)

Next, we use the spectral polynomial algorithm to find an expression for the distribution function of \((\alpha, T)\) in a few steps.

1. By using the spectral polynomial algorithm, we obtain \(TP(\lambda) = P(\lambda)S(\lambda)\), where \(P(\lambda)\) is an \(m \times m\) matrix with unit row sums and \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)\). (See Theorem 2 and the ensuing discussion.)

2. By using the spectral polynomial algorithm again, we obtain \(S(\mu, m)P_1 = P_1 S(\lambda)\), where \(P_1\) is a matrix with unit row sums. By Lemma 1, similarly to the proof of Theorem 3, it can be shown that \(P_1\) is invertible. Then the equation \(S(\mu, m)P_1 = P_1 S(\lambda)\) can be written as \(P_1^{-1} S(\mu, m) = S(\lambda) P_1^{-1}\).

3. Combining the above results, we obtain \(TP(\lambda) \cdot P_1^{-1} = P(\lambda)P_1^{-1} S(\mu, m)\) and \(P(\lambda)P_1^{-1} e = e\). It is easy to check that \(p_0, S(\mu, m)\) is an ME-representation of \((\alpha, T)\), where \(p_0, S(\mu, m)\) is an ME-representation of \((\alpha, T)\), where \(\beta = \alpha P(\lambda)P_1^{-1}\) (see the proof of Proposition 2.1 in He and Zhang 2006b).

**Lemma 2.** The distribution function \(F(t)\) of \((\alpha, T)\) can be obtained as (note \(m_0 = 0\))

\[ F(t) = 1 - \sum_{k=1}^{K} \left( \sum_{i=0}^{m_k-1} \frac{t^{i_k}}{i_k!} \left( \sum_{j=m_k+1}^{m_k+i+1} \beta_j \right) \right) \exp[-\mu_k t], \quad t \geq 0. \]
He and Zhang: An Algorithm for Computing Minimal Coxian Representations
183

Proof. By definitions and the structure of \( S(\mu, m) \), the distribution function \( F(t) \) of \((\alpha, T)\) can be obtained as

\begin{equation}
F(t) = 1 - \alpha \exp[Tt]e
= 1 - \beta \exp[S(\mu, m)t]e
= 1 - \sum_{k=1}^{K} \beta_{m_1+\cdots+m_{k-1}+1} \beta_{m_1+\cdots+m_{k-1}+2} \cdots, \tag{4}
\end{equation}

which leads to (4) by routine calculations. □

The \( PH \)-generator \( T \) may have non-real eigenvalues when the Laplace-Stieltjes transform of \((\alpha, T)\) has only real poles. However, because the Laplace-Stieltjes transform of the \( PH \)-distribution \((\alpha, T)\) has only real poles, the coefficient of \( \exp[-\mu_k t] \) in (4) must be zero if \( -\mu_k \) is not a real number. Without loss of generality and to simplify notation, we assume that all \( K \) eigenvalues \( \{-\mu_1, -\mu_2, \ldots, -\mu_K\} \) in (4) are real. Since \( T \) is a \( PH \)-generator, all its eigenvalues must have negative real parts (Minc 1988, Seneta 1973). Thus, \( -\mu_1, -\mu_2, \ldots, -\mu_K \) are negative real numbers and, consequently, \( S(\mu, m) \) is a \( PH \)-generator. Further, if \( \beta_{m_1+\cdots+m_{k-1}+1} = 0, 1 \leq i \leq m_k \), we remove all items associated with \( -\mu_k \). Thus, we assume that at least one of \( \{\beta_{m_1+\cdots+m_{k-1}+i} : 1 \leq i \leq m_k\} \) is not zero, for \( k = 1, 2, \ldots, K \). If \( \beta_{m_1+\cdots+m_k} = 0 \), then we choose the largest \( i \) such that \( \beta_{m_1+\cdots+m_{k-1}+i} \neq 0, 1 \leq i \leq m_k \), and redefine \( m_k \) to \( i \) to ensure \( \beta_{m_1+\cdots+m_k} \neq 0 \).

\[ N_m = m_1 + m_2 + \cdots + m_K, \tag{5} \]

which is the algebraic degree of the corresponding \( PH \)-distribution (O’Cinneide 1993).

It is readily seen that \( N_m \leq m \). (4) indicates that \( F(t) \) is associated with \( K \) Jordan blocks corresponding to distinct real eigenvalues \( \{-\mu_1, -\mu_2, \ldots, -\mu_K\} \) of \( T \). Next, we use the expression of \( F(t) \) given in (4) to prove that the algebraic degree \( N_m \) is a lower bound on the triangular order of the \( PH \)-distribution \((\alpha, T)\). It is well known that the triangular order and \( PH \)-order of a \( PH \)-distribution are as large as the algebraic degree of the \( PH \)-distribution (O’Cinneide 1993). We provide the following proof for completeness.

Lemma 3. Consider a \( PH \)-distribution with a \( PH \)-representation \((\alpha, T)\) of order \( m \). We assume \( \mu_1 > \mu_2 > \cdots > \mu_K \) and \( \beta_{m_1+\cdots+m_k} \neq 0 \) for \( k = 1, 2, \ldots, K \). Then any ME-representation of the \( PH \)-distribution \((\alpha, T)\) must have at least \( N_m \) phases. Consequently, the triangular order and the \( PH \)-order of \((\alpha, T)\) are as large or equal to \( N_m \).

Proof. By (4), since \( \min\{\mu_1, \ldots, \mu_K\} > \mu_K \), we have

\[ \lim_{t \to \infty} \frac{F(t)}{\mu_1^{m_1+\cdots+m_{k-1}} \mu_k^{m_k} \exp[-\mu_k t]} = \beta_{m_1+\cdots+m_k+1}. \tag{6} \]

Similarly to (6), for \( 1 \leq k \leq K-1 \), we have

\[ \lim_{t \to \infty} \frac{F(t) - \sum_{i=0}^{m_k-1} \beta_i t^{m_k-1}}{\mu_k^{m_k} \exp[-\mu_k t]} = \beta_{m_1+\cdots+m_{k-1}+1}. \]

Since \( \beta_{m_1+\cdots+m_k} \neq 0 \) for \( 1 \leq k \leq K \), it is easy to see that any expression of the distribution function of \((\alpha, T)\) must be equivalent to the one given in (4). Thus, the Laplace-Stieltjes transform of any representation of \( F(t) \) must have the poles \( -\mu_1, -\mu_2, \ldots, -\mu_K \) with multiplicities \( m_1, m_2, \ldots, m_K \), respectively. Therefore, any ME-representation of \((\alpha, T)\) must have at least \( N_m \) phases. Consequently, the triangular order and the \( PH \)-order of \((\alpha, T)\) are larger than or equal to \( N_m \). □

According to Lemma 3, \( (\beta, S(\mu, m), e) \) is a minimal ME-representation of \((\alpha, T)\) and the \( 3 \)-tuple \( (\beta, \mu, m) \) is invariant to the probability distribution \((\alpha, T)\).

4. A Set of Necessary and Sufficient Conditions

In this section, we use \( (\beta, \mu, m) \) defined in §3 to find a set of necessary and sufficient conditions for a Coxian representation of \((\alpha, T)\) to be minimal. We begin with the construction of an ME-representation with an ordered Coxian generator from \((\beta, \mu, m)\). Let \( n = (n_1, n_2, \ldots, n_{N_m}) \) with \( n_i = \mu_k \) for \( m_1 + m_2 + \cdots + m_{k-1} + 1 \leq j \leq m_1 + m_2 + \cdots + m_k \) and \( 1 \leq k \leq K \).

Since \( \mu_1 > \mu_2 > \cdots > \mu_K \), the elements of \( n \) are in nonincreasing order. By using the spectral polynomial algorithm, we can find \( P(n) \) satisfying \( S(\mu, m)P(n) = P(\eta)S(\eta) \) and \( P(\eta)e = e \). Since \( S(\mu, m) \) is lower triangular, by Corollary 2 in O’Cinneide (1989) (also see Theorem 4.3 in He and Zhang 2006b), \( P(\eta) \) is nonnegative so it is a stochastic matrix. Because all elements of \( \eta \) are positive, the determinant of \( S(\eta)e, (S(\eta))^2e, \ldots, (S(\eta))^{N_m}e \)

\[ \begin{pmatrix}
\eta_1 & -\eta_1^2 & \eta_1^3 & \cdots & (-1)^{N_m-1}\eta_1^{N_m} \\
0 & \eta_2 & -\eta_1^2 & \eta_1^3 & \cdots \\
0 & 0 & \eta_2 & \eta_1^2 & \cdots \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
0 & \cdots & \cdots & \cdots & 0 & \eta_{N_m-1} \eta_{N_m} \eta_{N_m} \\
0 & \cdots & \cdots & \cdots & 0 & \eta_1 \eta_2 \eta_3 \cdots \eta_{N_m}
\end{pmatrix}
\]

is nonzero. Thus, the vectors \( S(\eta)e, (S(\eta))^2e, \ldots, (S(\eta))^{N_m}e \) are independent. Therefore, by Theorem 1...
The relationship between $\gamma$, $S(x, \mathbf{e})$ defined above, we have $\gamma_j = h_{1,j}/(x_1x_2\cdots x_j)$, $1 \leq j \leq n$.

Proof. Since $\gamma = \beta P(x)$, we have $\gamma_j = \beta P_j(x) = 1 \leq j \leq n$. The results are obtained by (7) and (8) and routine calculations.

Lemma 5. The ME-representation $(\gamma, S(x, \mathbf{e})$ defined above is an ordered Coxian representation of $(\alpha, T)$ if and only if

1. $x_1 \geq x_2 \geq \cdots \geq x_n > 0$;
2. $\eta$ is part of $x$; and
3. $h_{1,j} \geq 0$, $1 \leq j \leq n$.

Proof. If $(\gamma, S(x, \mathbf{e})$ is an ordered Coxian representation of $(\alpha, T)$, we must have $x_1 \geq x_2 \geq \cdots \geq x_n > 0$. Thus, condition (1) holds. By Lemma 2, $\eta$ must be part of $x$, which implies condition (2). Since $\gamma \geq 0$, by Lemma 4, condition (3) holds. This proves the necessity of the conditions.

On the other hand, suppose that conditions (1), (2), and (3) hold. Condition (1) implies that $S(x)$ is a PH-generator. Condition (2) implies $S(x, \mathbf{e})$ is an equivalent ordered Coxian representation of $(\alpha, T)$, This proves the sufficiency of the conditions.

Theorem 4 in O’Cinneide (1993) provides a necessary and sufficient condition for a PH-representation with a triangular PH-generator to be minimal. Compared to that condition, the conditions given in Lemma 5 are easier to check numerically. In fact, the conditions in Lemma 5 lead to an algorithm for computing a minimal Coxian representation for any Coxian distribution. Lemma 5 establishes a relationship between the Coxian representation of a Coxian distribution and the derivatives of its distribution function at zero. Now, we are ready to state and prove the main result of this section.

Theorem 4. Let $n^*$ be the minimal $n$ such that conditions (1), (2), and (3) in Lemma 5 are satisfied. Then $n^*$ is the triangular order of $(\alpha, T)$. The number $n^*$ is finite and is at least $N_m$ (defined in (5)). The corresponding representation $(\gamma, S(x, \mathbf{e})$ is a minimal ordered Coxian representation of $(\alpha, T)$.

Proof. The first conclusion is obvious from Lemma 5. By Lemma 3, we must have $n^* \geq N_m$. According to Theorem 1, any PH-distribution with only real poles has a triangular representation or, equivalently, an ordered Coxian representation. Therefore, $n^*$ is finite.

It is clear from the proofs of the above lemmas that Theorem 4 is valid for any Coxian distribution with an ME-representation $(\alpha, T, \mathbf{e})$, where $\alpha$ may not be nonnegative. Consequently, the algorithm developed in §5 can be used for computing minimal ordered Coxian representations for Coxian distributions with ME-representations $(\alpha, T, \mathbf{e})$ (see Example 1, §6).
5. A Nonlinear-Programming Approach

Based on Lemma 5 and Theorem 4, a nonlinear system can be developed for computing a minimal ordered Coxian representation for \((\alpha, T)\). First, we prove a property for minimal ordered Coxian representations.

**Lemma 6.** For a minimal ordered Coxian representation \((\gamma, S(x))\) of order \(n^*\), we must have \(x_j \geq \eta_{N_{\alpha}}\), \(1 \leq j \leq n^*\).

**Proof.** Note that \(x_1 \geq x_2 \geq \cdots \geq x_n\). If there exists \(x_i\) such that \(x_j < \eta_{N_{\alpha}}\), then \(x_n\) is the smallest one with that property. Suppose that \(x_{n^*-1} > x_{n^*}\). By routine calculations, we can obtain

\[
\gamma \exp[S(x)t]e = \sum_{j=1}^{n^*-1} c_j(t) \exp[-x_j t] + \gamma_n \exp[-x_{n^*} t],
\]

where \(\{c_j(t), 1 \leq j \leq n^*\}\) are some polynomials of a finite order. By the proof of Lemma 3, we must have \(\gamma_n = 0\). Thus, the ordered Coxian representations \((\gamma_1, \gamma_2, \ldots, \gamma_{n^*}, S(x_1, x_2, \ldots, x_{n^*}))\) and \((\gamma, S(x))\) represent the same probability distribution. This further implies that \((\gamma, S(x))\) is not a minimal Coxian representation, which gives a contradiction. Therefore, for a minimal Coxian representation \((\gamma, S(x))\), we must have \(x_j \geq \eta_{N_{\alpha}}\) for \(1 \leq j \leq n^*\). The case \(x_{n^*-1} = x_{n^*}\) can be proved similarly. \(\square\)

Based on Lemma 5, Theorem 4, and Lemma 6, a Coxian representation can be constructed from any feasible solution to the following nonlinear system (NLS):

\[
\begin{align*}
(1) & \quad x_1 \geq x_2 \geq \cdots \geq x_n \geq \eta_{N_{\alpha}}, \quad x_1 \geq \eta_1; \\
(2) & \quad (\eta_1, \eta_2, \ldots, \eta_{N_{\alpha}}) \text{ is part of } (x_1, x_2, \ldots, x_n); \\
(3) & \quad h_{i,j} \geq 0, \quad 1 \leq j \leq n; \\
(4) & \quad (8) \text{ holds.}
\end{align*}
\]

Next, we discuss how to find a feasible solution to NLS (9). First, we can replace constraint (1) in (9) by \(x_1 \geq \eta_1\), and \(x_i \geq \eta_{N_{\alpha}}, 2 \leq i \leq n\). The solution to the modified nonlinear system (if exists) can be used to construct a Coxian representation, which may not be an ordered Coxian representation. By Property 5.1 in He and Zhang (2006b), we can use the spectral polynomial algorithm to find an ordered Coxian representation of the same order from that solution. Second, we rewrite (8) as the following recursive equations:

\[
\begin{align*}
 h_{i,1} &= F^{(0)}(0), \quad 1 \leq i \leq n; \\
 h_{i,j} &= x_{j-1} h_{i,j-1} + h_{i+j-1,j-1}, \\
 & \quad 3 \leq j + i \leq n + 1, \quad j \geq 2, \quad i \geq 1.
\end{align*}
\]

Equation (10) can be obtained in a straightforward manner from (8). Equalities in (10) are polynomials of degree 2 in \(\{x_1, x_2, \ldots, x_n\}\) and \(\{h_{i,j}, h_{2,j}, \ldots, h_{n+1,j}\}\), which is significantly smaller than that of (8). This may bring efficiency in solving the nonlinear system.

Let \(H = \{h_{i,j}, 2 \leq i + j \leq n + 1\}\). Combining Lemmas 5 and 6, Theorem 4, the above observations, and (10), we introduce the following nonlinear program (NLP) to find an ordered Coxian representation \((\gamma, S(x))\) for \((\alpha, T)\) and a given \(n (\geq N_{\alpha})\):

\[
\begin{align*}
\min & \quad \sum_{i=1}^{n} x_i \\
\text{s.t.} & \quad (1) \quad x_1 \geq \eta_1, \quad x_i \geq \eta_{N_{\alpha}}, \quad \text{for } 2 \leq i \leq n; \\
& \quad (2) \quad (\eta_1, \eta_2, \ldots, \eta_{N_{\alpha}}) \text{ is part of } (x_1, x_2, \ldots, x_n); \\
& \quad (3) \quad h_{i,j} \geq 0, \quad 1 \leq j \leq n; \\
& \quad (4) \quad (10) \text{ holds.}
\end{align*}
\]

If NLP (11) has a solution, that solution corresponds to a Coxian representation of order \(n\) for \((\alpha, T)\) (Lemma 5). By using the spectral polynomial algorithm, we can find an ordered Coxian representation of order \(n\) from that solution. Otherwise, the triangular order of \((\alpha, T)\) is larger than \(n\).

The derivative \(F^{(0)}(0)\) can be extremely large or small for large or even moderate \(j\), which can cause problems in computation. One suggestion to solve the problem is to rescale the time, i.e., to scale \(T \rightarrow \delta T\) where \(\delta\) is a positive number. For instance, we can choose \(\delta^{-1} = \max_{i \leq i \leq N_{\alpha}} |\lambda_i|\). Numerical examples demonstrate that such a change of scale makes the algorithm more stable.

From numerical experiments, we learned that choosing an appropriate initial search point is important for solving NLP (11) efficiently. An initial search point we use is \(x = (\lambda, \ldots, \lambda, \eta_1, \eta_2, \ldots, \eta_{N_{\alpha}})\), where \(\lambda > \eta_1\).

It is easy to see that the (minimal) ordered Coxian representation of \((\alpha, T)\) may not be unique. In fact, any feasible solution to (11) corresponds to an ordered Coxian representation of \((\alpha, T)\). Also, the optimal solution to (11) can be different if a different objective function is used. In (11), the objective function is a simple linear function. Such a linear function is chosen to make it easier to solve (11). The following property supports the nonuniqueness of minimal ordered Coxian representation.

**Lemma 7.** Suppose that

\[
(\gamma(0), S(x(0))) \quad \text{and} \quad (\gamma(1), S(x(1)))
\]

are two ordered Coxian representations of order \(n\) for \((\alpha, T)\), where \(x(0)\) and \(x(1)\) are identical except for one element (i.e., \(x_i(0) = x_i(1), 1 \leq i \leq n\) and \(i \neq i_0\)). Then for any convex combination \(x(\lambda) = \lambda x(1) + (1 - \lambda) x(0)\) with \(0 \leq \lambda \leq 1\), there is an ordered Coxian representation \((\gamma(\lambda), S(x(\lambda)))\) of order \(n\) for \((\alpha, T)\).
Proof. Because all elements of \(x(0)\) and \(x(1)\) are positive, all elements of \(x(\lambda)\) are positive. Since both \(x(0)\) and \(x(1)\) are ordered sequences, \(x(\lambda)\) is an ordered sequence as well. By (8), \([h_{ij}(0), 2 \leq i + j \leq n + 1]\) and \([h_{ij}(1), 2 \leq i + j \leq n + 1]\) are functions of the elements of \(x\). Define \([h_{ij}(0), 2 \leq i + j \leq n + 1]\) by (10) for \(x(0)\), \([h_{ij}(1), 2 \leq i + j \leq n + 1]\) for \(x(1)\), and \([h_{ij}(\lambda), 2 \leq i + j \leq n + 1]\) for \(x(\lambda)\). Since \(x_0(\lambda) = x_1(\lambda), 1 \leq i \leq n\), and \(x_0(\lambda) = x_1(\lambda) + (1 - \lambda) \cdot h_{i0}(0), 2 \leq i + j \leq n + 1\). Therefore, \([h_{ij}(\lambda), 1 \leq j \leq n]\) are nonnegative. Let \(\gamma(\lambda) = (\gamma_1(\lambda), \gamma_2(\lambda), \ldots, \gamma_n(\lambda))\) is part of \(x(\lambda)\). Then \(\gamma(\lambda), x(\lambda)\) satisfies the conditions in Lemma 5. Therefore, \((\gamma(\lambda), S(x(\lambda)))\) is an ordered Coxian representation of \((\alpha, T)\).

According to Aldous and Shepp (1987), for any \(PH\)-distribution \((\alpha, T)\) of order \(m\), we must have \(cv(\alpha, T) \geq 1/m\), where \(cv(\alpha, T)\) is the coefficient of variation of \((\alpha, T)\). That implies that \(m \geq 1/cv(\alpha, T) = (\alpha T^{-1} e)^2/[(\alpha T^{-1} e)^2 - (\alpha T^{-1} e)^2]\). Let \(N^*_m = \max[N_m, 1/cv(\alpha, T)]\). Apparently, \(N^*\) is a lower bound of the triangular order of \((\alpha, T)\).

Finally, we introduce an algorithm for computing a minimal ordered Coxian representation of \((\alpha, T)\). For this purpose, one needs to solve (11) for \(n = N^*, N^* + 1, \ldots, n\), until a solution is found. Theorem 4 ensures that the smallest \(n\) for which (11) has a solution is the triangular order of \((\alpha, T)\). We summarize the computational steps as follows.

**Minimal Coxian Representation Algorithm.** Assume that the Laplace-Stieltjes transform of \(PH\)-distribution \((\alpha, T)\) has only real poles.

**Step 1.** Find the spectrum of \(T\). Use the spectral polynomial algorithm to compute \((\beta, S(\mu, m))\) and \(N_m\) (defined in (5)). Remove all eigenvalues whose corresponding elements in \(\beta\) are all zero.

**Step 2.** Use the spectral polynomial algorithm to compute \(P(\eta)\) and \([\beta P(\eta), S(\eta)]\) from \((\beta, S(\mu, m))\). If \(\beta P(\eta)\) is nonnegative, \((\beta P(\eta), S(\eta))\) is a minimal ordered Coxian representation of order \(N_m\) for \((\alpha, T)\). Otherwise, compute \(N^*\) and let \(n = N^*\).

**Step 3.** Calculate \(F(0, 0) = -\alpha T^1 e, 1 \leq j \leq n + 1\). Solve (11). If there is no solution to (11), set \(n = n + 1\) and go to Step 3. Otherwise, use Lemma 4 to compute \(\gamma\) from the optimal solution. Then \((\gamma, S(x))\) is a minimal Coxian representation of order \(n\) for \((\alpha, T)\). If \((\gamma, S(x))\) is not an ordered Coxian representation (i.e., the elements in \(x\) are not in nonincreasing order), use the spectral polynomial algorithm to compute a minimal ordered Coxian representation from \((\gamma, S(x))\).

According to Cumani (1982) and He and Zhang (2006b), if \(T\) is triangular or symmetric, the search for the triangular order is between \(N^*\) and \(m\). In general, by Theorem 4.1 in O’Cinneide (1991), the search process will be terminated after a finite number of iterations. However, there is no known upper bound for the search process. Thus, finding an upper bound of the triangular order is useful for the above algorithm and is an interesting issue for future research.

### 6. Three Numerical Examples

In this section, we demonstrate the effectiveness of the algorithm developed in §§3, 4, and 5, show some geometric properties associated with ordered Coxian representations, and explain why finding a minimal ordered Coxian representation is not straightforward. Note that all distributions under consideration have a zero mass at \(t = 0\).

**Example 1.** We show that the triangular order of a \(PH\)-representation whose Laplace-Stieltjes transform has three poles can be arbitrarily large, as pointed out by O’Cinneide (1991). We consider a \(PH\)-generator \(T\) of order three given by

\[
T = \begin{pmatrix}
-7 & 0 & 0.3 \\
3 & -4 & 0 \\
0 & 3.5 & -4
\end{pmatrix}.
\]

The eigenvalues of \(T\) are \([-\lambda_1 = -6.4933, -\lambda_2 = -5.4056, -\lambda_3 = -3.1012]\). Let \(\alpha_1, \alpha_2, \alpha_3\) be the corresponding eigenvectors normalized by \(\alpha_1 \cdot e = \alpha_2 \cdot e = \alpha_3 \cdot e = 1\). Then \((|\alpha_1, T, e|, 1 \leq i \leq 3)\) represent three exponential distributions with parameters \((\lambda_1, \lambda_2, \lambda_3)\), respectively. Similarly to Example 3.2 in He and Zhang (2006a) (also see examples in Dehon and Latouche 1982), we construct a \(PH\)-invariant polytope \(conv[q_1, q_{12}, q_{123}]\) from \((\alpha_1, \alpha_2, \alpha_3)\) as follows:

\[
q_1 = \alpha_1;
q_{12} = \frac{\lambda_2}{\lambda_2 - \lambda_1 - \alpha_1} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \alpha_2;
q_{123} = \frac{\lambda_3 \lambda_2}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)} \alpha_1 + \frac{\lambda_1 \lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)} \alpha_2 + \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)} \alpha_3.
\]

Let \(Q\) be a \(3 \times 3\) matrix with \(q_1\) as its first row, \(q_{12}\) as its second row, and \(q_{123}\) as its third row. It can be shown that \(QT = S(\lambda)Q\). Since \(S(\lambda)\) is a \(PH\)-generator, we say that the polytope \(conv[q_1, q_{12}, q_{123}]\) is \(PH\)-invariant under \(T\).

A polytope is a convex set with a finite number of extreme points. See Rockafellar (1970) for more about convex sets and polytopes. The concept of \(PH\)-invariant polytopes was introduced in O’Cinneide (1991), who discusses more about \(PH\)-invariant polytopes.
According to Dehon and Latouche (1982), conv \([q, q_{12}, q_{123}]\) contains all \(\alpha\) corresponding to probability distributions \((\alpha, T, e)\) with a bidiagonal PH-representation of order three or smaller. Denote by \(\Omega_3\) the set of all vectors \(\alpha\) with a unit sum corresponding to probability functions that are affine combinations of exponential distributions \((\alpha_i, T, e), 1 \leq i \leq 3\). Denote by \(e\), the unit row vector with all elements being zero except the \(i\)-th element, which is one, \(1 \leq i \leq 3\). By definition, all probability vectors are in the polytope \(\text{conv}\{e_1, e_2, e_3\}\). We have \(\text{conv}\{e_1, e_2, e_3\} \subseteq \Omega_3\) and \(\text{conv}\{q, q_{12}, q_{123}\} \subseteq \Omega_3\). In Figure 1, the probability vector polytope \(\text{conv}\{e_1, e_2, e_3\}\), the polytope \(\text{conv}\{q, q_{12}, q_{123}\}\), and the convex set \(\Omega_3\) are plotted.

For any PH-representation \((\alpha, T)\) with \(T\) given in (12), if \(\alpha\) is in \(\text{conv}\{q, q_{12}, q_{123}\}\), \((\alpha, T)\) has an ordered Coxian representation of order three or smaller. Otherwise, \((\alpha, T, e)\) does not have an ordered Coxian representation of order three. He and Zhang (2006b) show that PH-representation \((\alpha, T)\) of order three always has an ordered Coxian representation of order four or smaller. For instance, the triangular order of PH-distribution \((e_1, T)\) is four, three for \((e_2, T)\), and three for \((e_3, T)\). Note that the triangular order of \((e_1, T)\) is larger than its PH-order (which is three). If \(\alpha\) is not in \(\text{conv}\{q, q_{12}, q_{123}\}\) nor in \(\text{conv}\{e_1, e_2, e_3\}\), then the triangular order of \((\alpha, T, e)\) is at least four and can be arbitrarily large.

Next, we consider the following vectors \(\alpha\) in \(\Omega_3\) and use the nonlinear program developed in §§3, 4, and 5 to find the triangular order and a minimal ordered Coxian representation of \((\alpha, T, e)\). If \(\alpha = \alpha_i\), \(i = 1, 2, \) or 3, we have \(n^* = 1\). If \(\alpha = y\alpha_1 + (1 - y)\alpha_3\), for \(0 < y < 1\) and \(i = 1\) or 2, we have \(n^* = 2\). If \(\alpha = y\alpha_1 + x\alpha_3 + (1 - y - x)\alpha_3\) for \(0 < y, z, y + z < 1\), we have \(n^* = 3\). Now, we choose \(\alpha\) outside of \(\text{conv}\{q, q_{12}, q_{123}\}\) as follows: \(\alpha = 0.5(1, 0, 0) + 0.5q_{123} + \delta(1, -1, 0)\), where \(\delta\) is to be determined. For the following \(\alpha\), we use NLP (11) to find the triangular order \(n^*\) of \((\alpha, T, e)\):

\[
\delta = 0: \quad \alpha_4 = (0.4714, -0.1458, 0.6744), \quad n^* = 4;
\]
\[
\delta = 0.7: \quad \alpha_5 = (1.1714, -0.8458, 0.6744), \quad n^* = 5;
\]
\[
\delta = 0.8: \quad \alpha_6 = (1.2714, -0.9458, 0.6744), \quad n^* = 6;
\]
\[
\delta = 0.9: \quad \alpha_7 = (1.3714, -1.0458, 0.6744), \quad n^* = 8;
\]
\[
\delta = 1.0: \quad \alpha_8 = (1.4714, -1.1458, 0.6744), \quad n^* = 12;
\]
\[
\delta = 1.1: \quad \alpha_9 = (1.5714, -1.2458, 0.6744), \quad n^* = 31;
\]
\[
\delta = 1.2: \quad \alpha_{10} = (1.6714, -1.3458, 0.6744), \quad n^* = \text{nonexistent}.
\]

The vectors \(\{\alpha_i, 4 \leq i \leq 10\}\) are plotted in Figure 1 as well. It is clear that when \(\alpha\) approaches the boundary of \(\Omega_3\), the triangular order of \((\alpha, T, e)\) can be very large. When \(\delta = 1.2\), the corresponding \(\alpha_{10}\) is outside of \(\Omega_3\). Thus \((\alpha_{10}, T, e)\) does not represent a probability distribution. Numerical experimentations demonstrate that the minimal Coxian representation algorithm is efficient if \(\alpha\) is not close to the boundary of \(\Omega_3\).

For Example 1, the original PH-generator \(T\) is not an ordered Coxian generator. In the next two examples, we consider cases for which \(T\) is an ordered Coxian generator. As will be shown, finding a minimal Coxian representation can be equally complicated even if the original representation is an ordered Coxian representation.

**Example 2.** We consider a PH-generator \(S(\lambda)\) of order four given by

\[
S(\lambda) = \begin{pmatrix}
-10 & 0 & 0 & 0 \\
5 & -5 & 0 & 0 \\
0 & 1.5 & -1.5 & 0 \\
0 & 0 & 1 & -1
\end{pmatrix}.
\]

Eigenvalues of \(S(\lambda)\) are \(-\lambda_1 = -10, -\lambda_2 = -5, -\lambda_3 = -1.5, -\lambda_4 = -1\), and the corresponding eigenvectors are \(\alpha_1 = (1, 0, 0, 0), \alpha_2 = (0.5, 0.5, 0, 0), \alpha_3 = (0.15, 0.255, 0.595, 0), \alpha_4 = (0.1, 0.18, 0.48, 0.24)\), respectively. All eigenvectors are normalized to have a unit sum. According to He and Zhang (2006b), \(\text{conv}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\) is a PH-invariant polytope under \(S(\lambda)\). Again, by using the method developed in Dehon and Latouche (1982), the PH-invariant polytope \(\text{conv}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\) can be expanded to a PH-invariant polytope \(\text{conv}\{q_1, q_{12}, q_{123}, q_{1234}\}\) as follows:

\[
q_1 = \alpha_1; \quad q_{12} = \frac{\lambda_1}{\lambda_2 - \lambda_1} \alpha_1 + \frac{\lambda_2}{\lambda_1 - \lambda_2} \alpha_2;
\]
\[
q_{123} = \frac{\lambda_2 \lambda_3}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)} \alpha_1 + \frac{\lambda_1 \lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)} \alpha_2
\]
\[
+ \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)} \alpha_3;
\]
\[
q_{1234} = \frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 - \lambda_4)(\lambda_2 - \lambda_4)(\lambda_3 - \lambda_4)} \alpha_1 + \frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)(\lambda_4 - \lambda_3)} \alpha_2
\]
\[
+ \frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)(\lambda_4 - \lambda_2)} \alpha_3
\]
\[
+ \frac{\lambda_1 \lambda_2 \lambda_3}{(\lambda_1 - \lambda_4)(\lambda_2 - \lambda_4)(\lambda_3 - \lambda_4)} \alpha_4.
\]
\[ q_{1234} = \frac{\lambda_4 \lambda_3 \lambda_2}{(\lambda_4 - \lambda_1)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} \alpha_1 + \frac{\lambda_4 \lambda_2 \lambda_1}{(\lambda_4 - \lambda_2)(\lambda_2 - \lambda_1)} \alpha_2 + \frac{\lambda_4 \lambda_3 \lambda_1}{(\lambda_4 - \lambda_3)(\lambda_3 - \lambda_1)} \alpha_3 + \frac{\lambda_2 \lambda_3 \lambda_1}{(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} \alpha_4. \]

For this example, because \( S(\lambda) \) is itself an ordered Coxian generator, it can be verified that \( q_1 = e_1, q_{12} = e_2, q_{123} = e_4 \), and \( q_{1234} = e_4 \). According to Dehon and Latouche (1982), the simplex \( \Gamma_{1,2,3,4} = \text{conv}\{e_1, e_2, e_3, e_4\} \) consists of all probability vectors \( \alpha \) such that \((\alpha, S(\lambda), e) \) is a representation of an ordered Coxian distribution of order four or smaller. Let \( \Omega_{1,2,3,4} \) be the set of all vectors \( \alpha \) with a unit sum such that \((\alpha, S(\lambda), e) \) is an ME-representation of a distribution.

We consider four sub-affine sets generated by \( \{\alpha_1, \alpha_2, \alpha_3\}, \{\alpha_1, \alpha_3, \alpha_4\}, \{\alpha_2, \alpha_3, \alpha_4\}, \{\alpha_2, \alpha_4\} \), and \( \{\alpha_1, \alpha_3, \alpha_4\} \), respectively. Denote the sub-affine sets as \( \text{aff}\{\alpha_1, \alpha_2, \alpha_3\}, \text{aff}\{\alpha_1, \alpha_3, \alpha_4\}, \text{aff}\{\alpha_2, \alpha_3, \alpha_4\}, \text{aff}\{\alpha_2, \alpha_4\}, \text{aff}\{\alpha_1, \alpha_3, \alpha_4\} \). By Dehon and Latouche (1982), the set of all probability distributions in each sub-affine set can be identified, which are denoted by \( \Omega_{1,2,3}, \Omega_{1,3,4}, \Omega_{1,2,4}, \Omega_{1,2,3,4}, \Omega_{2,3,4,5} \) respectively. These subsets of probability distributions and polytopes are plotted in Figure 2. Similarly to the expansion from \( \text{conv}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} \) to \( \text{conv}\{q_1, q_{12}, q_{123}, q_{1234}\} \), we can expand the PH-invariant polytopes \( \text{conv}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} \) to \( \Gamma_{1,2,3} = \text{conv}\{q_1, q_{12}, q_{123}\} \), \( \Gamma_{1,2,4} = \text{conv}\{q_1, q_{12}, q_{123}, q_{1234}\} \), \( \Gamma_{1,3,4} = \text{conv}\{q_1, q_{12}, q_{123}, q_{1234}\} \), and \( \Gamma_{2,3,4} = \text{conv}\{q_2, q_{23}, q_{234}\} \), respectively. All the ordered Coxian representations of order three with eigenvalues chosen from \( \{-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4\} \) in these PH-invariant polytopes.

Intuitively, for a probability vector \( \alpha \) in some subspace, it is possible that \((\alpha, S(\lambda), e) \) has an ordered Coxian representation of lower order. In fact, if a probability vector \( \alpha \) is in \( \text{aff}\{\alpha_1, \alpha_2, \alpha_3\} \), \( \text{aff}\{\alpha_1, \alpha_3, \alpha_4\} \), or \( \text{aff}\{\alpha_2, \alpha_3, \alpha_4\} \), then \((\alpha, S(\lambda), e) \) has an ordered Coxian representation of order three or smaller. The reason is that \( \Omega_{1,2,3} \cap \Gamma_{1,2,3,4} = \Gamma_{1,2,3}, \Omega_{1,2,4} \cap \Gamma_{1,2,3,4} = \Gamma_{1,2,4}, \Omega_{1,3,4} \cap \Omega_{1,2,3,4} = \Omega_{1,3,4,5} \) as shown in Figure 2.

Figure 2. \( \Omega_{1,2,3}, \Omega_{1,2,4}, \Omega_{1,3,4}, \Omega_{2,3,4}, \) and PH-Invariant Polytopes for Example 2

However, if a probability vector \( \alpha \) is in \( \text{aff}\{\alpha_2, \alpha_4\} \), it is possible that \( F(t) \) of \((\alpha, S(\lambda), e) \) has an ordered Coxian representation of order four whose Laplace-Stieltjes transform has poles \(-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4\) and the triangular order of \( F(t) \) is four, i.e., \( F(t) \in \Gamma_{1,2,3,4} \) but \( F(t) \notin \Gamma_{2,3,4} \). The reason is that \( \Gamma_{2,3,4} \subset \Omega_{2,3,4} \cap \Gamma_{1,2,3,4} \) and \( \Gamma_{2,3,4} \neq \Omega_{2,3,4} \cap \Gamma_{1,2,3,4} \) as shown in Figure 2. In fact, \( \Omega_{2,3,4} \cap \Gamma_{1,2,3,4} = \text{conv}\{q_2, q_{23}, q_{234}, q_0\} \), where \( q_0 \) is in the set \( \text{conv}\{q_1, q_{1234}\} \) (the point “*” in Figure 2). Any ordered Coxian distribution in the interior of the set \( \text{conv}\{q_2, q_{234}, q_0\} \) has a triangular order four.

In summary, for any Coxian distribution in \( \Gamma_{1,2,3}, \Gamma_{1,3,4} \), and \( \Gamma_{1,2,4} \), the number of poles of its Laplace-Stieltjes transform equals its triangular order. For a Coxian distribution in \( \Gamma_{2,3,4} \), it is possible that the number of the poles of its Laplace-Stieltjes transform is smaller than its triangular order.

Example 3. We consider a Coxian distribution \((\alpha, S(\lambda)) \) with \( m = 5 \), \( \alpha = (0.0787, 0.0778, 0.0184, 0.1174, 0.7076) \), and

\[
S(\lambda) = \begin{pmatrix} -20 & 0 & 0 & 0 & 0 \\ 10 & -10 & 0 & 0 & 0 \\ 0 & 7.5 & -7.5 & 0 & 0 \\ 0 & 0 & 5 & -5 & 0 \\ 0 & 0 & 0 & 1.5 & -1.5 \end{pmatrix}, \quad (13)
\]

where \( \lambda = (20, 10, 7.5, 5, 1.5) \). It can be shown that the PH-distribution \((\alpha, S(\lambda)) \) has an ME-representation \((\beta, S((7.5, 5, 1.5)), e) \) of order three with poles \((-7.5, -5, -1.5) \) and \( \beta = (0.21, -0.11, 0.9) \). By Lemma 2, the triangular order of \((\alpha, S(\lambda)) \) must be at least four.

To see if \((\alpha, S(\lambda)) \) has an ordered Coxian representation of order four, we consider ordered Coxian generators with eigenvalues \((-10, -7.5, -5, -1.5) \) and \((-20, -7.5, -5, -1.5) \), where \(-10 \) and \(-20 \) are the eigenvalues of the original PH-generator given in (13). Using the spectral polynomial algorithm, it can be shown that \((\alpha, S(\lambda)) \) is equivalent to ME-representations \((\beta_1, S((10, 7.5, 5, 1.5)), e) \) and \((\beta_2, S((20, 7.5, 5, 1.5)), e) \), \( \beta_1 = (0.1575, -0.0025, 0.0800, 0.7650) \) and \( \beta_2 = (0.0787, 0.1038, -0.0150, 0.8235) \). Unfortunately, none of them is a PH-representation. Therefore, we need to try a value other than \(-10 \) and \(-20 \). By considering \(-15 \), i.e., \( \lambda = (15, 7.5, 5, 1.5) \), we find \( \beta_2 = (0.1050, \ldots) \).
0.0683, 0.0167, 0.8100), which implies that \((\mathbf{b}_\lambda, S((15, 7.5, 5, 1.5)))\) is an alternative \(PH\)-representation of \((\alpha, S(\lambda))\). Since \((\mathbf{b}_\lambda, S((15, 7.5, 5, 1.5)))\) is an ordered Coxian representation, the triangular order of \((\alpha, S(\lambda))\) is four. Furthermore, using Lemma 7, we find that we can replace \(-15\) by any number between \(-10.88\) and \(-18.00\) to generate an alternative ordered Coxian representation of \((\alpha, S(\lambda))\).

An interesting observation is that the eigenvalue \(-15\) is not an eigenvalue of the original generator \(S(\lambda)\). Thus, to find a minimal ordered Coxian representation for \((\alpha, S(\lambda))\), we must find the eigenvalue \(-15\) (or a number between \(-10.88\) and \(-18.00\)), which is not straightforward.

7. Discussion on Future Research

An interesting issue for future research is to explore the structure of NLP (11) to develop more efficient ways to solve it. Also, we would like to find an upper bound on the triangular order, which makes it possible to combine all the nonlinear programs into a single nonlinear program to find a minimal Coxian representation.

Another area of future research is to extend the spectral polynomial method to parameter estimation and fitting of \(PH\)-distributions, which are of great practical significance. Possible results in this area may complement to those in Asmussen et al. (1996), Bobbio et al. (2002, 2004), Feldmann and Whitt (1998), and Johnson and Taaffe (1989; 1990a, b).

Acknowledgments

The authors thank an anonymous referee for detailed and constructive comments and suggestions on an earlier version of this paper. This research project was financially supported by the National Sciences and Engineering Research Council (NSERC) of Canada and the Chinese Academy of Sciences (People’s Republic of China).

References


