# The Chebyshev-Markov Inequalities

## Xinyun Wang and Martin Haenggi

#### I. INTRODUCTION

In many applications, such as computed tomography (CT) image reconstructions [1], distributions of bounded support need to be reconstructed from moment sequences. Without loss of generality, we assume the support of the distributions is [0, 1]. The problem can be formulated as follows. Let  $[n] \triangleq \{1, 2, ..., n\}$  and  $[n]_0 \triangleq \{0\} \cup [n]$ . Given a finite sequence  $(m_k)_{k=0}^n$ ,  $n \in \mathbb{N}$ , find an F that solves

$$\int_0^1 x^k \, dF(x) = m_k, \quad \forall k \in [n]_0,\tag{1}$$

where F is right-continuous and increasing with  $F(0^-) = 0$  and F(1) = 1, i.e., F is a cumulative distribution function (cdf). This problem is known as the truncated Hausdorff moment problem (THMP) [2]. Let  $\mathcal{F}_n$  denote the set of all possible F that solve (1). Assuming such solutions exist, it is natural to consider the sharpest bounds of F at the point of interest, i.e.,  $\inf_{F \in \mathcal{F}_n} F(x_0)$ and  $\sup_{F \in \mathcal{F}_n} F(x_0)$  for any  $x_0 \in [0, 1]$ .

The problem that asks for the sharpest bounds of F was first stated by Chebyshev [3] and later solved by Markov [4]. Possé [5] simplified the original proof of Markov. Zelen [6] was the first one to state the sharpest bounds in a generalized form, i.e., he gave the formulation of the sharpest bounds not only in the bounded support case, but also in the unbounded support cases such as  $(-\infty, \infty)$ ,  $[0, \infty)$ , and  $(-\infty, 0]$ . The inequalities established by the sharpest bounds are called the Chebyshev-Markov (CM) inequalities [6]. This report is mainly based on the results of [2], [5], [6], and we only consider the case of bounded support.

X. Wang and M. Haenggi are with the Dept. of Electrical Engineering, University of Notre Dame, Indiana, USA. Email: {xwang54,mhaenggi}@nd.edu. This work was supported in part by the US National Science Foundation through Grant 2007498.

### II. THE CM INEQUALITIES

Markov [4] provided a method to obtain the infimum and supremum

$$\inf_{F \in \mathcal{F}_n} F(x_0), \quad \sup_{F \in \mathcal{F}_n} F(x_0)$$
(2)

for any  $x_0 \in [0, 1]$ . The most important step of the method is the construction of a discrete distribution in  $\mathcal{F}_n$  where the maximum mass is concentrated at  $x_0$ . Let  $p_0$  denote the maximum mass that is possible to be concentrated at  $x_0$  and  $F^*$  denote the discrete distribution where the maximum mass is concentrated at  $x_0$ , i.e.,  $F^*(x_0) - F^*(x_0^-) = p_0$ . Then,  $\inf_{F \in \mathcal{F}_n} F(x_0) = F^*(x_0^-)$ and  $\sup_{F \in \mathcal{F}_n} F(x_0) = F^*(x_0)$ . In the following, we recall the details of the method to construct  $F^*$ . As for any discrete distribution, there are jump locations and jump heights (probability masses concentrated at the jumps). Suppose that  $F^*$  is constructed by jumps at  $x_i$  with heights  $p_i$ ,  $1 \le i \le v, v \in \mathbb{N}$ . If we know  $(x_i)_{i=1}^v$ , then  $(p_j)_{i=1}^v$  can be obtained by solving

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_v \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \dots & x_v^n \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_v \end{pmatrix} = \begin{pmatrix} m_0 \\ m_1 \\ \vdots \\ m_n \end{pmatrix}.$$
(3)

To find  $(x_i)_{i=1}^v$  and further construct  $F^*$ , we recall the following definition and lemma.

**Definition 1** (Orthogonal polynomials w.r.t. measures [7]). An orthogonal polynomial of degree m w.r.t. a measure dF, associated with the moment sequence  $(m_k)_{k=0}^{2m-1}$ , is given by<sup>1</sup>

$$\begin{vmatrix} m_0 & m_1 & \dots & m_m \\ m_1 & m_2 & \dots & m_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{m-1} & m_m & \dots & m_{2m-1} \\ 1 & x & \dots & x^m \end{vmatrix}$$
(4)

**Lemma 1** ([2], [6]). Let  $t_r, u_r, v_r, w_r$  be the orthogonal polynomials of degree  $r \in \mathbb{N}$  w.r.t. the measures dF, ydF, (1-y)dF and y(1-y)dF associated with the moment sequences  $(m_k)_{k=0}^{2r-1}$ ,  $(m_{k+1})_{k=0}^{2r-1}$ ,  $(m_k - m_{k+1})_{k=0}^{2r-1}$  and  $(m_{k+1} - m_{k+2})_{k=0}^{2r-1}$ , respectively. For a moment sequence  $(m_k)_{k=0}^n$ ,

<sup>&</sup>lt;sup>1</sup>We ignore the arbitrary constant factor since we are only interested in the roots.

1) if  $x_0 \in (0,1)$  is distinct from the roots of  $u_m, v_m$  (for n = 2m) and  $t_m, w_{m-1}$  (for n = 2m - 1), let the polynomial q of degree l be defined as

$$(y-x_0)\omega(y), \quad n=2m, \ u_m(x_0)v_m(x_0) > 0,$$
(5)

$$q(y) \triangleq \begin{cases} (y-1)y(y-x_0)\omega(y), & n = 2m, \ u_m(x_0)v_m(x_0) < 0, \end{cases}$$
(6)

$$y(y - x_0)\omega(y), \quad n = 2m - 1, \ t_m(x_0)w_{m-1}(x_0) > 0, \tag{7}$$

$$(y-1)(y-x_0)\omega(y), \quad n=2m-1, \ t_m(x_0)w_{m-1}(x_0)<0,$$
(8)

where l = m + 2 for (6), l = m + 1 for the others, and  $\omega(y)$  is the orthogonal polynomial of degree m w.r.t. the measure  $(y - x_0)dF$  associated with the moment sequence  $(m_{k+1} - x_0m_k)_{k=0}^{2m-1}$  for (5) and of degree m-1 w.r.t. the measures  $(y-1)y(y-x_0)dF$ ,  $y(y-x_0)dF$ and  $(y-1)(y-x_0)dF$  associated with the moment sequences  $(m_{k+3} - (1+x_0)m_{k+2} + x_0m_{k+1})_{k=0}^{2m-3}$ ,  $(m_{k+2} - x_0m_{k+1})_{k=0}^{2m-3}$  and  $(m_{k+2} - (1+x_0)m_{k+1} + x_0m_k)_{k=0}^{2m-3}$  for (6), (7) and (8), respectively. The roots of the polynomial q of degree l are in [0, 1], they coincide with the jumps  $(x_i)_{i=1}^{v}$ , and  $x_0$  is one of them.

- if x<sub>0</sub> ∈ (0,1) is a root of u<sub>m</sub> or v<sub>m</sub> (for n = 2m) and t<sub>m</sub> or w<sub>m-1</sub> (for n = 2m − 1), it is clear that the roots of the corresponding orthogonal polynomial are all in [0,1],<sup>2</sup> and the roots plus 0 and/or 1 coincide with the jumps (x<sub>i</sub>)<sup>v</sup><sub>i=1</sub>.
- 3) if  $x_0 = 0$ , let the polynomial q of degree l be defined as

$$\begin{cases}
y\omega(y), \quad u_m(x_0)v_m(x_0) > 0, \\
(9)
\end{cases}$$

$$q(y) \triangleq \left\{ (y-1)y\omega(y), \quad u_m(x_0)v_m(x_0) < 0 \text{ or } t_m(x_0)w_{m-1}(x_0) < 0, \quad (10) \right\}$$

$$(y\omega(y), t_m(x_0)w_{m-1}(x_0) > 0,$$
 (11)

where  $m = \lfloor \frac{n+1}{2} \rfloor$ , l = m for (11), l = m + 1 for the others, and  $\omega(y)$  is the orthogonal polynomial of degree m w.r.t. the measure ydF associated with the moment sequence  $(m_{k+1})_{k=0}^{2m-1}$  for (9) and of degree m-1 w.r.t. the measures (y-1)ydF and ydF associated with the moment sequences  $(m_{k+2} - m_{k+1})_{k=0}^{2m-3}$  and  $(m_{k+1})_{k=0}^{2m-3}$  for (10) and (11), respectively. The roots of the polynomial q of degree l are in [0, 1], they coincide with the jumps  $(x_i)_{i=1}^{v}$ , and  $x_0$  is one of them.

3

<sup>&</sup>lt;sup>2</sup>The roots of  $t_m$  are different from those of  $u_m$ . The roots of  $v_m$  are different from those of  $w_{m-1}$ .

4) if  $x_0 = 1$ , let the polynomial q of degree l be defined as

$$(y-1)\omega(y), \quad u_m(x_0)v_m(x_0) > 0,$$
(12)

$$q(y) \triangleq \begin{cases} (y-1)y\omega(y), & u_m(x_0)v_m(x_0) < 0 \text{ or } t_m(x_0)w_{m-1}(x_0) > 0, \end{cases}$$
(13)

$$\left( (y-1)\omega(y), \quad t_m(x_0)w_{m-1}(x_0) < 0, \right.$$
(14)

where  $m = \lfloor \frac{n+1}{2} \rfloor$ , l = m for (14), l = m + 1 for the others, and  $\omega(y)$  is the orthogonal polynomial of degree m w.r.t. the measure (y-1)dF associated with the moment sequence  $(m_{k+1}-m_k)_{k=0}^{2m-1}$  for (12) and of degree m-1 w.r.t. the measures (y-1)ydF and (y-1)dFassociated with the moment sequences  $(m_{k+2}-m_{k+1})_{k=0}^{2m-3}$  and  $(m_{k+1}-m_k)_{k=0}^{2m-3}$  for (13) and (14), respectively. The roots of the polynomial q of degree l are in [0, 1], they coincide with the jumps  $(x_i)_{i=1}^{v}$ , and  $x_0$  is one of them.

**Theorem 1** ([2]). For the truncated Hausdorff moment problem with a moment sequence  $(m_k)_{k=0}^n$ , for all  $x_0 \in [0, 1]$ , let  $F^*$  denote the discrete distribution constructed by jump locations  $(x_i)_{i=1}^v$ obtained in Lemma 1 and jump heights  $(p_i)_{i=1}^v$  obtained by solving (3). Then

$$\inf_{F \in \mathcal{F}_n} F(x_0) = F^*(x_0^-) = \sum_{j: x_j < x_0} p_j,$$
(15)

$$\sup_{F \in \mathcal{F}_n} F(x_0) = F^*(x_0) = \sum_{j: x_j \le x_0} p_j.$$
(16)

The inequalities established by the infima and suprema obtained in Theorem 1 are the CM inequalities.

In the following, we provide two examples with n = 1 and n = 2, which prove the well-known Markov's inequality and Chebyshev's inequality, respectively.

**Example 1** (n = 1). For n = 1, we have  $t_1(x) = x - m_1$ ,  $w_0(x) = 1$  and  $\omega(y) = 1$ . If  $x_0 < m_1$ , the roots of q are  $x_0$  and 1, thus  $0 \le F(x_0) \le \frac{1-m_1}{1-x_0}$ ; if  $x_0 > m_1$ , the roots of q are 0 and  $x_0$ , thus  $1 - \frac{m_1}{x_0} \le F(x_0) \le 1$ ; if  $x_0 = m_1$ ,  $0 \le F(x_0) \le 1$ . The lower bound is equivalent to Markov's inequality.

**Example 2** (n = 2). For n = 2, we have  $u_1(x) = m_1 x - m_2$  and  $v_1(x) = (m_0 - m_1)x - (m_1 - m_2)$ . Consider the case where all the Hankel determinants are positive. Then  $m_0m_2 > m_1^2$  and thus  $\frac{m_1 - m_2}{m_0 - m_1} < \frac{m_2}{m_1}$ . If  $x_0 < \frac{m_1 - m_2}{m_0 - m_1}$ , then  $\omega(y) = (m_1 - x_0m_0)y - (m_2 - x_0m_1)$ , the roots of q are  $x_0$  and  $\frac{m_2 - x_0m_1}{m_1 - x_0m_0}$ , and  $x_0 < \frac{m_2 - x_0m_1}{m_1 - x_0m_0}$ . Thus  $0 \le F(x_0) \le \frac{m_2 - m_1^2}{(x_0 - m_1)^2 + m_2 - m_1^2}$ . If  $x_0 > \frac{m_2}{m_1}$ , then  $\omega(y) = (m_1 - x_0 m_0)y - (m_2 - x_0 m_1), \text{ the roots of } q \text{ are } x_0 \text{ and } \frac{m_2 - x_0 m_1}{m_1 - x_0 m_0}, \text{ and } x_0 > \frac{m_2 - x_0 m_1}{m_1 - x_0 m_0}.$ Thus  $\frac{(x_0 - m_1)^2}{(x_0 - m_1)^2 + m_2 - m_1^2} \leq F(x_0) \leq 1.$  If  $\frac{m_1 - m_2}{m_0 - m_1} < x_0 < \frac{m_2}{m_1}$ , then  $\omega(y) = 1$ , the roots of q are  $0, x_0$ and 1, and  $1 - m_1 + \frac{m_2 - m_1}{x_0} \leq F(x_0) \leq 1 - m_1 - \frac{m_2 - m_1}{1 - x_0}.$  If  $x_0 = \frac{m_2}{m_1}$ , the roots are 0 and  $x_0$ , and  $1 - \frac{m_1^2}{m_2} \leq F(x_0) \leq 1.$  If  $x_0 = \frac{m_1 - m_2}{m_0 - m_1}$ , the roots are  $x_0$  and 1, and  $0 \leq F(x_0) \leq \frac{(m_0 - m_1)^2}{m_0 - 2m_1 + m_2}.$ We can prove Chebyshev's inequality in this way.<sup>3</sup>

#### REFERENCES

- T. J. Wang and T. Sze, "The image moment method for the limited range CT image reconstruction and pattern recognition," *Pattern Recognition*, vol. 34, no. 11, pp. 2145–2154, 2001.
- [2] J. Shohat and J. Tamarkin, *The Problem of Moments*, ser. Mathematical surveys and monographs. American Mathematical Society, 1950.
- [3] P. L. Chebyshev, Sur les valeurs limites des intégrales. Imprimerie de Gauthier-Villars, 1874.
- [4] A. Markoff, "Démonstration de certaines inégalités de M. Tchébychef," *Mathematische Annalen*, vol. 24, no. 2, pp. 172–180, 1884.
- [5] K. A. Possé, Sur quelques applications des fractions continues algébriques. L'Académie Impériale des Sciences, 1886.
- [6] M. Zelen, "Bounds on a distribution function that are functions of moments to order four," Journal of Research of the National Bureau of Standards, vol. 53, no. 6, pp. 377–81, 1954.
- [7] G. Szeg, Orthogonal polynomials. American Mathematical Society, 1939, vol. 23.

<sup>&</sup>lt;sup>3</sup>Zelen [6] has also proved Chebyshev's inequality as a special case of n = 4. Chebyshev's inequality can also be proved by Markov's inequality.