Songklanakarin J. Sci. Technol.
43 (3), 744-751, May - Jun. 2021

http://www.sjst.psu.ac.th

Original Article

A method for approximate missing data from data error measured with l^{∞} norm

Benjawan Rodjanadid[1] and Kannika Khompungson[2*]

[1] *School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Mueang, 30000 Thailand*

[2] *Division of Mathematics, School of Science, University of Phayao, Phayao, Mueang, 56000 Thailand*

Received: 3 November 2019; Revised: 2 April 2020; Accepted: 20 May 2020

Abstract

We briefly review some recent work on hypercircle inequality for partially corrupted data when the data error is measured with l^{∞} norm. The aim of this paper is to present the method for approximate missing data in the use of midpoint algorithm and discuss some numerical experiments in the hardy space which is well known in a reproducing kernel Hilbert space (RKHS).

Keywords: hypercircle inequality, convex optimization, reproducing kernel Hilbert space

1. Introduction

The extension of hypercircle inequality to partially corrupted data was proposed by Khompurngson and Novaprateep (2015). That is, the data set contains both accurate and inaccurate data. Specifically, the material on this subject has been applied to kernel-based machine learning problem when data set contains both accurate and inaccurate data, see e.g. Khompurngson and Suantai (2016). The aim of this paper is to present the method to approximate missing data in the use of midpoint algorithm and discuss some numerical experiments in hardy space Hardy space of square integrable function on unit circle. Specifically, we consider the data error is measured with l^{∞} norm. In this section, we describe hypercircle inequality and its potential relevance to kernel-based learning.

Let $X=\{x_i : i\in N_n\}$ denote the set of linearly independent vectors in Hilbert space H over the real numbers with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$ where we use the notation $N_n=\{1,2,...,n\}$. The operator $Q:H\rightarrow\mathbb{R}^n$ is defined by for $x\in H$

$$Qx=(\langle x,x_j\rangle : j\in N_n).$$

Recalling that the adjoint map is $Q^T:\mathbb{R}^n\rightarrow H$ defined by the property that for all $a\in\mathbb{R}^n$ and $x\in H$

$$\langle Q^T(a),x\rangle=(a,Qx).$$

*Corresponding author
Email address: kannika.kh@up.ac.th

Therefore, it follows that Q^T is given at $a=(a_j : j\in N_n)\in\mathbb{R}^n$

$$Q^T(a)=\sum_{j\in\mathbb{N}_n} a_j x_j.$$

Definition 1.1 For $d=\{d_i : i\in N_n\}$, the set of $x\in H$ such that $Qx=d$ and $\|x\|\le 1$ is called a *hypercircle,* $\mathrm{H}(d)$. That is, we have

$$\mathrm{H}(d)=\{x:\|x\|\le 1, Qx=d\},$$

see Khompurngson and Micchelli (2011).
Moreover, it is well-known that there is a unique vector $x(d)\in M$ such that

$$x(d)=\arg\min\{\|x\|: x\in H, Qx=d\} \qquad (1.1)$$

where *M* is the n-dimensional subspace of *H* spanned by the vectors in *X*, see Davis (1975). Moreover, the vector

$$x(d)=Q^T G^{-1}d \quad \text{and} \quad \|x(d)\|=(d,G^{-1}d),$$

where the gram matrix *G* is defined by

$$G=\begin{bmatrix}\langle x_1,x_1\rangle & \langle x_1,x_2\rangle & \dots & \langle x_1,x_n\rangle\\ \langle x_2,x_1\rangle & \langle x_2,x_2\rangle & \dots & \langle x_2,x_n\rangle\\ \vdots & \vdots & \dots & \vdots\\ \langle x_n,x_1\rangle & \langle x_n,x_2\rangle & \dots & \langle x_n,x_n\rangle\end{bmatrix}.$$

Therefore, the hypercircle inequality states the following way:

Theorem 1.2 If $x\in\mathrm{H}(d)$ and $x_0\in H$ then

$$\left|\langle x(d),x_0\rangle-\langle x,x_0\rangle\right|\le dist(x_0,M)\sqrt{1-\|x(d)\|^2}$$

where $dist(x_0,M):=\min\{\|x_0-y\|: y\in M\}$. Moreover, there are elements $x_\pm(d)\in\mathrm{H}(d)$ for which equality above holds and the vector $x_\pm(d)$ are given by the formula

$$x_\pm(d)=\pm\frac{x_0-Q^T c_\pm}{\left\|x_0-Q^T c_\pm\right\|} \quad \text{and} \quad c_\pm:=G^{-1}\Big(Qx_0\mp\frac{dist(x_0,M)}{\sqrt{1-\|x(d)\|^2}}\Big).$$

The detailed proofs appear in Khompurnson and Micchelli (2011). Moreover, let us point that the best value to estimate $\langle x,x_0\rangle$ is the midpoint, $\langle x(d),x_0\rangle$, of this interval $I(x_0)=\{\langle x,x_0\rangle : x\in\mathrm{H}(d)\}$.

Next let us specialize the material above to the problem of learning the value of a function in hardy space Hardy space of square integrable function on unit circle. Let $\mathrm{H}^2(\Delta)$ be the set of all functions analytic in the unit disc Δ with norm $\|f\|=\sup_{0<r<1}\Big(\frac{1}{2\pi}\int_0^{2\pi}|f(re^{i\theta})|^2\,d\theta\Big)^{\frac{1}{2}}$. The reproducing kernel for $\mathrm{H}^2(\Delta)$ is given by

$$K(s,t)=\frac{1}{1-st}, s,t\in\Delta,$$

see Duren (2000). Let $T=\{t_1,t_2,\dots,t_n\}$ be points of increasing order in (-1,1). Consequently, we have a finite set of linearly independent elements $\{K_{t_i} : i\in N_n\}$ in *H* where

$$K_{t_i}(t)=\frac{1}{1-t_i t}, i\in N_n$$

and $t\in\Delta$. Thus, the vectors $\{x_i : i\in N_n\}$ are identified with the function $\{K_{t_i} : i\in N_n\}$ Therefore, the Gram matrix of the $\{K_{t_i} : i\in N_n\}$ is given by

$$G(t_1,t_2,...,t_n)=(K(t_i,t_j):i,j\in N_n)$$

According to Cauchy determinant, see Davis (1975), which state that for any $\{t_i : i\in N_n\},\{s_i : i\in N_n\}$, we obtain that

$$\det\left(\frac{1}{1-s_it_j}\right)_{i,j\in\mathbb{N}_n}=\frac{\prod_{1\le j<i\le n}(t_j-t_i)(s_j-s_i)}{\prod_{i,j\in\mathbb{N}_n}(1-t_is_j)} \tag{9}$$

From this formula we have

$$\det G(t_1,...,t_n)=\frac{\prod_{1\le i<j\le n}(t_i-t_j)^2}{\prod_{i,j\in\mathbb{N}_n}(1-t_it_j)} \tag{10}$$

In our case, let us recall the function B , which is called a Blashke product of degree n with zeros in the set *T*, see Duren (2000). For any $t_0\in(-1,1)$ and $t_0\notin T=\{t_1,t_2,...,t_n\}$ we obtain that

$$dist(K_{t_0},span\{K_{t_i}:j\in N_n\})=\frac{|B(t_0)|}{\sqrt{1-t_0^2}},$$

where B is the rational function defined at $t\in\mathbb{C}\setminus\{t_i^{-1}:j\in N_n\}$ by

$$B(t):=\prod_{j\in\mathbb{N}_n}\frac{t-t_j}{1-tt_j}$$

and the vector x_0 appearing previously is identified with the function K_{t_0}, see Khompurnson and Micchelli (2011). These formulas allow us to give explicit hypercircle bounds for $\langle K_{t_0},f\rangle=f(t_0)$.

Proposition 1.3 If $f\in H^2(\triangle)$ with $f(t_i)=d_i, i\in N_n$ then for any $t\in(-1,1)$ we have that

$$f_d(t)-\frac{|B(t)|}{\sqrt{1-t^2}}\sqrt{1-\theta}\le f(t)\le f_d(t)+\frac{|B(t)|}{\sqrt{1-t^2}}\sqrt{1-\theta}\text{ , where}$$

$$f_d(t)=\sum_{j\in\mathbb{N}_n}\frac{B(t)d_j}{B'(t)(t-t_j)}\text{ and }\theta:=\sum_{j\in\mathbb{N}_n}\frac{d_jd_l}{B'(t_j)B'(t_l)(1-t_jt_l)}\ ,$$

see Khompurnson and Micchelli (2011).

The organization of this paper is presented as follows. In the next section, we review a basic fact about hypercircle inequality for partially corrupted data with l^∞ norm. The material of this is described in the generality of an arbitrary Hilbert space (not just a reproducing kernel Hilbert space). In section 3, we discuss some numerical experiments of learning the value of a function from data error measured with l^∞ norm Hardy space of square integrable function on unit circle.

2. Hypercircle inequality for data error measured with l^∞ norm

In this section, we briefly review Hypercircle inequality for partially corrupted data (Hipcd) measured with l^∞ norm and refer the paper of Khompurnson and Novaprateep (2015) for the proof of (Hipdc).

Definition 2.1 Let $E_\infty=\{e:e\in\mathbb{R}^n, e_I=0, \||e_J\||_\infty\le\varepsilon\}$ and ε is some positive number. For each $d=(d_j:j\in N_n)\in\mathbb{R}^n$, the set of $x\in H$ such that $Qx-d\in E_\infty$ and $\|x\|\le 1$ is called a *partial hyperellipse* $\mathrm{H}(d\,|\,E_\infty)$. That is,

$$\mathrm{H}(d\,|\,E_\infty)=\{x:Qx-d\in E_\infty,\|x\|\le 1\}.$$

Given $x_0 \in H$ we want to estimate $\langle x, x_0 \rangle$ knowing that $\| x \| \leq 1$ and $\langle x, x_i \rangle = d_i$ for all $i \in I$ and the data error $\||(d - Qx)_J\||_\infty \leq \varepsilon$. That is, the exact data becomes $\langle x, x_i \rangle = d_i$ for all $i \in I$ and the data error, which means data that have the error from a variety of reason in practice, becomes $e = \langle x, x_j \rangle - d_j$ which has absolute value $\leq \varepsilon$ and $\varepsilon > 0$ is prescribed. Moreover, let us point out the relation between the definition of hypercircle and hyperellipse as follows.

$$\mathrm{H}(d \mid E_\infty) = \bigcup_{e \in E_\infty} \mathrm{H}(d + e)$$

In our case, we also see that

$$\mathrm{H}(d \mid E_\infty) = H(d_I) \cap \bigcap_{\substack{|\mathrm{e}_i| \leq \varepsilon \\ i \in J}} \mathrm{H}(d_i + \mathrm{e}_i)$$

In the same manner, we can see that the best estimator is the midpoint of the uncertainty interval

$$I(x_0, d \mid E_\infty) = \{\langle x, x_0 \rangle : x \in \mathrm{H}(d \mid E_\infty)\}.$$

According to our previous work, let us point out that the uncertainty interval

$I(x_0, d \mid E_\infty) = [m_-(x_0, d \mid E_\infty), m_+(x_0, d \mid E_\infty)]$ is closed and bounded interval on $\mathbb{R}$ where

$$m_+(x_0, d \mid E_\infty) := \max\{\langle x, x_0 \rangle : x \in H(d \mid E_\infty)\}$$

and $m_-(x_0, d \mid E_\infty) := \min\{\langle x, x_0 \rangle : x \in H(d \mid E_\infty)\}$ respectively. Consequently, the midpoint is given by

$$m(x_0, d \mid E_\infty) = \frac{m_+(x_0, d \mid E_\infty) - m_+(x_0, -d \mid E_\infty)}{2}$$

which follows from $m_-(x_0, d \mid E_\infty) = -m_+(x_0, -d \mid E_\infty)$. The proposition below is important to determine the best estimator in the partial hyperellipse $\mathrm{H}(d \mid E_\infty)$ for estimating $\langle x, x_0 \rangle$ when $x \in \mathrm{H}(d \mid E_\infty)$.

Theorem 2.2 If $\mathrm{H}(d \mid E_\infty) \neq \phi$ then there exist $x_\pm(d \mid E_\infty) \in \mathrm{H}(d \mid E_\infty)$ such that

$$m_\pm(x_0, d \mid E_\infty) = \langle x_\pm(d \mid E_\infty), x_0 \rangle.$$

Moreover, we know that there are the vectors $e_\pm \in E_\infty$ such that

$$m_\pm(x_0, d \mid E_\infty) = \langle x(d + e_\pm), x_0 \rangle \pm \mathrm{dist}(x_0, \mathrm{M}) \sqrt{1 - \| x(d + e_\pm) \|^2}$$

and the best estimator can be chosen on the line segment joining e_- and e_+.
That is, there is $e_0 = \lambda_0 e_+ + (1 - \lambda_0) e_-$ such that

$$x(d + e_0) = Q^T G^{-1}(d + e_0) \in M \quad \text{and} \quad m(x_0, d \mid E_\infty) = \langle x(d + e_0), x_0 \rangle,$$

where the value $\lambda_0 \in [0,1]$.

Proof. See Khompurnson and Novaprateep (2015).
Therefore, hypercircle inequality for partially-corrupted data error state below.

Theorem 2.3 If $x_0 \in H$ and $\mathrm{H}(d \mid E_\infty) \neq \phi$ then there is an $e_0 \in E_\infty$ such that for any $x \in \mathrm{H}(d \mid E_\infty)$

$$\left|\langle x(d+e_0), x_0\rangle - \langle x, x_0\rangle\right| \leq \frac{m_+(x_0, d \mid E_\infty) - m_+(x_0, -d \mid E_\infty)}{2},$$

Proof. See Khompurnson and Novaprateep (2015).
For this purpose, let us recall the duality formula for the right hand end point

$$m_+(x_0, d \mid E_\infty) = \max\{\langle x, x_0\rangle : x \in \mathrm{H}(d \mid E_\infty)\}$$

of the uncertainty interval. To this end, we define the function $V : \mathbb{R}^n \to \mathbb{R}^n$ as for each $c \in \mathbb{R}^n$ by

$$\mathrm{V}(c) := \| x_0 - Q^T c \| + \varepsilon \||| c_J \||_1 + (c, d).$$

Theorem 2.4 If $\mathrm{H}(d \mid E_\infty)$ contain more than one element then

$$m_+(x_0, d \mid E_\infty) = \min\{\mathrm{V(c)} : c \in \mathbb{R}^n\}. \tag{2.1}$$

Moreover, $c^* = \arg\min\{\mathrm{V(c)} : c \in \mathbb{R}^n\}$ is unique solution of (2.1).

Proof. See Khompurnson and Novaprateep (2015).

Our task is now to compute the solution of (2.1) explicit. For a fix $e \in \mathrm{H}(d \mid E_\infty)$, we define the following notation $Z(e) := \{j : j \in J, \mathrm{e}_j = 0\}$ and $\mathrm{U}(e) := \mathrm{J} \setminus \mathrm{Z(e)}$. Consequently, we use that notation $|z(e)|$ denoted the number element in $Z(e)$.

Theorem 2.5 If $x_0 \notin M$ and $\mathrm{H}(d \mid E_\infty)$ contain more than one element then we have that the following:

1. $\dfrac{x_0}{\| x_0 \|} \in \mathrm{H}(d \mid E_\infty)$ if and only if $m_+(x_0, d \mid E_\infty) = \| x_0 \|$

2. $x_+(d_I) \in \mathrm{H}(d \mid E_\infty)$ if and only if $|Z(\mathrm{c}^*)| = |\mathrm{J}|$ and $m_+(x_0, d \mid E_\infty) = \langle x(d_I), x_0\rangle + dist(x_0, M_I)\sqrt{1 - \| x(d_I) \|^2}$ where the vector $x_+(d_I) = \arg\max\{\langle x, x_0\rangle : x \in H(d_I)\}$.

3. If $x_+(d_{I \cup J_0}) \in \mathrm{H}(d \mid E_\infty)$ for some $J_0 \subseteq J$ then there exists unique $c \in \mathbb{R}^n$ with $|Z(c^*)| < |\mathrm{J}|$ such that the vector

$$x_+(d \mid E_\infty) := \frac{x_0 - Q^T c^*}{\| x_0 - Q^T c^* \|} = \arg\ \max\{\langle x, x_0\rangle : x \in H(d_I) \cap \bigcap_{\substack{|\mathrm{e}_i| \leq \varepsilon \\ i \in J_0}} \mathrm{H}(d_i + \mathrm{e}_i)\}.$$

Moreover, we have the following $m_+(x_0, d \mid E_\infty) = \min\{\| x_0 - Q^T_{I \cup J_0} c \| + \varepsilon \||| c_{J_0} \||_1 + (c, d_{I \cup J_0}) : c \in \mathbb{R}^{|I \cup J_0|}\}$.

4. If $\dfrac{x_0}{\| x_0 \|}, x_+(d_{I \cup J_0}) \notin \mathrm{H}(d \mid E_\infty)$ for each $J_0 \subseteq J$ then there exists unique $c^* \in \mathbb{R}^n$ with $|Z(c^*)| = 0$ such that the vector $x_+(d \mid E_\infty) := \dfrac{x_0 - Q^T c^*}{\| x_0 - Q^T c^* \|} \in \mathrm{H}(d \| \bar{E}_\infty)$

Moreover, the vector $x_+(d \mid E_\infty) := \arg\ \max\{\langle x, x_0\rangle : x \in \mathrm{H}(d \mid \bar{E}_\infty)\}$ and $\bar{E}_\infty := \{e : e \in E_\infty, |e_i| = \varepsilon, \forall i \in J\}$

Proof. See Khompurnson and Novaprateep (2015), Khompurngson and Nammanee (2019).

Theorem 2.6 If $x_0 := Q^T c^0 \in M$ and $\mathrm{H}(d \mid E_\infty)$ contain more than one element then $c^0 = \arg\min\{\mathrm{V(c)} : c \in \mathbb{R}^n\}$ if and only if $\mathrm{H}(d \mid E_\infty(x_0)) \neq \phi$.

Proof. We first assume that $c_0 = \arg\min\{\mathrm{V(c)} : c \in \mathbb{R}^n\}$. Consequently, we obtain that $m_+(x_0, d \mid E_\infty) = \varepsilon \||| \mathrm{c}^0_J \||_1 + (\mathrm{c}^0, d)$.

That is, there exist $\hat{x} \in \mathrm{H}(d \mid E_\infty)$ such that $\langle Q^T(c^0), \hat{x} \rangle = \varepsilon ||| c_J^0 |||_1 + (c^0, d)$

It implies $\left(c^0, Q\hat{x} - d\right) = \varepsilon ||| c_J^0 |||_1$ which means that $\hat{x} \in \mathrm{H}(d \mid E_\infty(x_0))$.

Conversely, we assume that $\mathrm{H}(d \mid E_\infty(x_0)) \neq \phi$ and then obtain that

$$
\begin{aligned}
\mathrm{m}_+(x_0, d \mid E_\infty) &= \min\{||| x_0 - Q^T c \,|| + \varepsilon ||| c_J |||_1 + (c, d) : c \in \mathbb{R}^n \} \\
&\leq \varepsilon ||| c_J^0 |||_1 + (c^0, d)
\end{aligned}
$$

Let $x \in \mathrm{H}(d \mid E_\infty(x_0)) \subseteq \mathrm{H}(d \mid E_\infty)$ and we see that

$$
\begin{aligned}
\mathrm{m}_+(x_0, d \mid E_\infty) &\geq \langle Q^T c^0, x \rangle \\
&= \left(c^0, Qx - d\right) + (c^0, d) \\
&= \varepsilon ||| c_J^0 |||_1 + (c^0, d)
\end{aligned}
$$

Therefore, we obtain that

$$\mathrm{m}_+(x_0, d \mid E_\infty) = \varepsilon ||| c_J^0 |||_1 + (c^0, d).$$

3. Application

In this section, we continue to report further a computational experiment in the use of midpoint algorithm. Specifically, we let $\mathrm{H}^2(\Delta)$ be the set of all functions analytic in the unit disc Δ with norm $\| f \| = \sup_{0<r<1} \left(\frac{1}{2\pi} \int_0^{2\pi} | f(re^{i\theta}) |^2 \, d\theta\right)^{\frac{1}{2}}$.

The reproducing kernel for $\mathrm{H}^2(\Delta)$ is given by

$$K(s,t) = \frac{1}{1 - st}, s, t \in \Delta.$$

We choose a finite set of linear independent elements $\{K_{t_j} : j \in N_6\}$ in $\mathrm{H}^2(\Delta)$ where $t_1 = -0.6, t_2 = -0.4, t_3 = -0.2,$ $t_4 = 0.2, t_5 = 0.4$ and $t_6 = 0.6$. We choose the exact function is $g(t) = -0.15K_{0.5}(t) + 0.05K_{0.85}(t) - 0.25K_{-0.5}(t)$ and compute the vector $d = (g(t_j) : j \in N_6)$. In our computation, we choose $J = \{k, l\} \subseteq N_6$ and then $E_\infty = \{e : e \in \mathbb{R}^2, e_I = 0, ||| e_J |||_\infty \leq \varepsilon\}$. Given $0 = t_0$ we wish to estimate $f(0) = \langle f, K_{t_0} \rangle$ when our data is given by $\| f \| \leq \delta$, $f(t_j) = d_j$ for all $i \in I$ and the data error $| f(t_k) - d_k | \leq \varepsilon$ and $| f(t_l) - d_l | \leq \varepsilon$. Using material form Theorem 2.4. we approximate the missing data which is defined as the value of data that is not stored for a variable.

In our computation, we extend partially hyperellipse to the case that the unit ball B is replaced by δB where δ is positive numbers. The computational steps for all the examples are organized in the following way. Within this propose, let us provide the important cases of the existence of the minimum of the convex function of δV. To this end, let us recall the following vector

$$c^* = \arg\min\{\delta V(c) : c \in \mathbb{R}^n \}$$

Where the function δV defined for $C \in \mathbb{R}^n$ as

$$\delta V = \sqrt{1 - 2\sum_{i \in \mathbb{N}_n} c_i + \sum_{i,j \in \mathbb{N}_n} \frac{c_i c_j}{1 - t_i t_j}} + \varepsilon \| c_J \|_1 + \sum_{i,j \in \mathbb{N}_n} c_i d_i$$

Proposition 3.1 If $K_{t_0} \notin M$ and $\mathrm{H}(d \mid E_\infty)$ contain more than one element then we have that the following:

1. V achieve its minimum with $| Z(c^*) | = 3$ If $\delta\sqrt{1 - t_0^2} K_{t_0} \in \mathrm{H}(d \mid E_\infty)$ then

$$m_+(\mathrm{K}_{t_0}, d \mid E_\infty) = \frac{\delta}{\sqrt{1 - t_0^2}}.$$

2. V achieve its minimum with $|Z(c^*)|=2$ If $f_{d_I}^+ := \delta \frac{K_{t_0} - Q_I^T c_+}{\| K_{t_0} - Q_I^T c_+ \|} \in H(d \mid E_\infty)$ then

$$m_+(K_{t_0}, d \mid E_\infty) = \sum_{j\in I} \frac{B_I(t_0) d_j}{B_I'(t_0)(t_0 - t_j)} + \frac{|B_I(t_0)|}{\sqrt{1-t_0^2}} \sqrt{\delta^2 - \beta}$$

where $f_{d_I}(t) = \sum_{j\in I} \frac{B(t) d_j}{B'(t)(t-t_j)}$, $c_+ := G^{-1}(Q_I K_{t_0} - \frac{dist(K_{t_0}, M_I)}{\sqrt{\delta^2 - \| f_{d_I} \|^2}})$ and $\beta := \sum_{j\in I} \frac{d_j d_l}{B'(t_j) B'(t_l)(1-t_j t_l)}$.

3. V achieve its minimum with $|Z(c^*)|=1$ If $\delta\sqrt{1-t_0^2} K_{t_0}, f_{d_I}^+ \notin H(d \mid E_\infty)$ and $\delta \frac{K_{t_0} - Q^T_{I\cup\{j\}} b^*}{\| K_{t_0} - Q^T_{I\cup\{j\}} b^* \|} \in H(d \mid E_\infty)$ for some $j \in J = \{k, l\}$ then we have that

$$m_+(K_{t_0}, d \mid E_\infty) = \delta \left\langle \frac{K_{t_0} - Q^T_{I\cup\{j\}} b^*}{\| K_{t_0} - Q^T_{I\cup\{j\}} b^* \|}, K_{t_0} \right\rangle = \frac{\delta}{\sqrt{1 - 2\sum_{i\in I\cup\{j\}} b^*_i + \sum_{i,p\in I\cup\{j\}} \frac{b^*_i b^*_p}{1-t_i t_p}}} \left(1 - \sum_{i\in I\cup\{j\}} b^*_i\right)$$

and $b^* := G_{I\cup\{j\}}^{-1}(Q_{I\cup\{j\}} K_{t_0} - \frac{dist(K_{t_0}, M_{I\cup\{j\}})}{\sqrt{\delta^2 - \| f_{d|E_j} \|^2}})$

where $f_{d|E_j} := \arg\min\{ f : \| f \| \le \delta, f(t_i) = d_i, \forall i \in I, | f(t_j) - d_j | \le \varepsilon \}$.

4. V achieve its minimum with $|Z(c^*)|=0$ If $\delta\sqrt{1-t_0^2} K_{t_0}, f_{d_I}^+, \delta \frac{K_{t_0} - Q^T_{I\cup\{j\}} b^*}{\| K_{t_0} - Q^T_{I\cup\{j\}} b^* \|} \notin H(d \mid E_\infty)$ for each $j \in J = \{k, l\}$ then

$$m_+(K_{t_0}, d \mid E_\infty) = \delta \left\langle \frac{K_{t_0} - Q^T c^*}{\| K_{t_0} - Q^T c^* \|}, K_{t_0} \right\rangle = \frac{\delta}{\sqrt{1 - 2\sum_{i\in \mathbb{N}_n} c^*_i + \sum_{i,p\in \mathbb{N}_n} \frac{c^*_i c^*_p}{1-t_i t_p}}} \left(1 - \sum_{i\in \mathbb{N}_n} c^*_i\right)$$

and $c^* := G^{-1}(Q K_{t_0} - \frac{dist(K_{t_0}, M)}{\sqrt{\delta^2 - \| f_{d|\bar{E}} \|^2}})$

where $f_{d|\bar{E}} := \arg\min\{ f : \| f \| \le \delta, f(t_i) = d_i, \forall i \in I. f(t_j) = d_j \pm \varepsilon, \forall j \in J \}$.

For our computation, we consider the partially hyperellipse in the different way as follows. In the first case, we choose the data error $| f(t_k) - d_k | \le \varepsilon$ for $t_k = -0.2, 0.2$, $| f(t_k) - d_k | \le \varepsilon$ for $t_k = -0.4, 0.4$ and , $| f(t_k) - d_k | \le \varepsilon$ for $t_k = -0.6, 0.6$ in the second and third case, respectively. The results of this computation are shown in Table 1 and the exact value of $g(0) = 2$. Our computation indicates that the midpoint algorithm on the learning tasks provided, at least in this numerical experiment is close to the exact value of $g(0) = 2$.

Table 1. Optimal value, $m(t_0, d \mid E_\infty)$, for different of data error

$J = \{t_i, t_j\}$	$m(t_0, d \mid E_\infty)$
-0.2, 0.2	2.0106
-0.4, 0.4	2.0034
-0.6, 0.6	2.0053

4. Conclusions

In this paper, we have provide some basic fact about hypercircle inequality for partially corrupted data when data error is measured with l^{∞} norm. We provided some results of numerical experiments of learning the missing data in Hardy space. To obtain the best estimator in the general case, we continue in the same method as in proposition 3.1 to obtain the solution of *V*.

References

Davis, P. J. (1975). *Interpolation and approximation*. New York, NY: J Dover Publications.

Duren, P. (2000). *Theory of H^P space* (2nd ed.). New York, NY: J Dover Publications.

Khompurngson, K., & Nammanee, K. (2019), Hypercircle inequality for data error measured with l^{∞} norm. Manuscript submitted for publication.

Khompurngson, K., & Micchelli, C. A. (2011). Hide. *Jaen Journal on Approximation*, *3*, 87-115.

Khompurngson, K., & Novaprateep, B. (2015). Hypercircle inequality for partially corrupted data. *Annals of Functional Analysis, 6*, 95-108.

Khompurngson, K., & Suantai, S.(2016). Alternative approximation method for learning multiple feature. *Thai Journal of Mathematic*, *14*, 517-526.