Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 4, Issue 3, Article 57, 2003

NEW NORM TYPE INEQUALITIES FOR LINEAR MAPPINGS

SABUROU SAITOH

DEPARTMENT OF MATHEMATICS,
FACULTY OF ENGINEERING
GUNMA UNIVERSITY,
KIRYU 376-8515, JAPAN.
ssaitoh@math.sci.gunma-u.ac.jp

Received 09 December, 2002; accepted 02 July, 2003
Communicated by S.S. Dragomir

ABSTRACT. In this paper, in connection with a basic formula by S. Smale and D. X. Zhou which is fundamental in the approximation error estimates in statistical learning theory, we give new norm type inequalities based on the general theory of reproducing kernels combined with linear mappings in the framework of Hilbert spaces. We shall give typical concrete examples which may be considered as new norm type inequalities and have physical meanings.

Key words and phrases: Learning theory, Convergence rate, Approximation, Reproducing kernel, Hilbert space, Linear mapping, Norm inequality.

2000 *Mathematics Subject Classification.* Primary 68T05, 30C40, 44A05, 35A22.

1. INTRODUCTION AND RESULTS

In statistical learning theory, reproducing kernel Hilbert spaces are a basic tool. See [1] for an excellent survey article and [4] for some recent very interesting results. In this paper, using a simple and general principle we shall show that we can obtain new norm type inequalities based on [3]. See [3] for the details in connection with learning theory. It seems that we can obtain new norm type inequalities based on a general principle which has physical meanings in linear systems. In order to show the results clearly, we shall first state our typical results.

For any fixed $q > 0$, let L_q^2 be the class of all square integrable functions on the positive real line $(0, \infty)$ with respect to the measure $t^{1-2q}dt$. Then, we consider the Laplace transform

$$(LF)(x) = \int_0^\infty F(t)e^{-xt}dt \qquad \text{for} \qquad x > 0$$

for $F \in L_q^2$.

ISSN (electronic): 1443-5756

139-02

Theorem 1.1. *For $LF = f$ and $LG = g$, we obtain the inequality*

$$\inf_{\|G\|_{L_q^2} \leq R} \frac{1}{\Gamma(2q+1)} \int_0^\infty (f'(x) - g'(x))^2 x^{2q+1} dx \leq \|F\|_{L_q^2}^2 \left(1 - \frac{R}{\|F\|_{L_q^2}}\right)^2,$$

for $\|F\|_{L_q^2} \geq R$.

We shall consider the Weierstrass transform, for any fixed $t > 0$

$$u_F(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} F(\xi) \exp\left\{-\frac{(x-\xi)^2}{4t}\right\} d\xi,$$

for $L_2(\mathbf{R}, d\xi)$ functions. Then,

Theorem 1.2. *For $u_F(x,t)$ and $u_G(x,t)$, we have the inequality*

$$\inf_{\|G\|_{L_2} \leq R} \|u_F(\cdot,t) - u_G(\cdot,t)\|_{L_2} \leq \|F\|_{L_2} \left(1 - \frac{R}{\|F\|_{L_2}}\right),$$

for $\|F\|_{L_2} \geq R$.

For any fixed $a > 0$ we shall consider the Hilbert space $\mathbf{F}_a$ consisting of entire functions $f(z)$ on $\mathbf{C}(z = x + iy)$ with finite norms

$$\|f\|_{\mathbf{F}_a}^2 = \frac{a^2}{\pi} \int \int_{\mathbf{C}} |f(z)|^2 \exp(-a^2|z|^2) dxdy.$$

Then we have

Theorem 1.3. *For any f and $g \in \mathbf{F}_a$, we have the inequality*

$$\inf_{\|g\|_{\mathbf{F}_a} \leq R} \int_{-\infty}^{\infty} |f(x) - g(x)|^2 \exp(-a^2x^2) dx \leq \frac{\sqrt{2\pi}}{a} \|f\|_{\mathbf{F}_a}^2 \left(1 - \frac{R}{\|f\|_{\mathbf{F}_a}}\right)^2,$$

for $\|f\|_{\mathbf{F}_a} \geq R$.

It was a pleasant suprise for the author that Professor Michael Plum was able to directly derive a proof of Theorem 1.2 in a Problems and Remarks session in the 8th General Inequalities Conference. In this paper, we will be able to give some general background for Theorem 1.2. Furthermore, we can obtain various other norm inequalities of type Theorem 1.2.

2. A General Approach and Proofs

The proofs of the theorems are based on a simple general principle, and on some deep and delicate parts which depend on case by case, arguments.

We need the general theory of reproducing kernels which is combined with linear mappings in the framework of Hilbert spaces in [2].

For any abstract set E and for any Hilbert (possibly finite–dimensional) space $\mathcal{H}$, we shall consider an $\mathcal{H}$–valued function $\mathbf{h}$ on E

$$\mathbf{h} : E \longrightarrow \mathcal{H} \tag{2.1}$$

and the linear mapping for $\mathcal{H}$

$$f(p) = (\mathbf{f}, \mathbf{h}(p))_{\mathcal{H}} \quad \text{for} \quad \mathbf{f} \in \mathcal{H} \tag{2.2}$$

into a linear space comprising of functions $\{f(p)\}$ on E. For this linear mapping, we consider the function $K(p,q)$ on $E \times E$ defined by

$$K(p,q) = (\mathbf{h}(q), \mathbf{h}(p))_{\mathcal{H}} \quad \text{on} \quad E \times E. \tag{2.3}$$

Then, $K(p,q)$ is a positive matrix on E; that is, for any finite points $\{p_j\}$ of E and for any complex numbers $\{C_j\}$,

$$\sum_j \sum_{j'} C_j \overline{C_{j'}} K(p_{j'}, p_j) \geq 0.$$

Then, by the fundamental theorem of Moore–Aronszajn, we have:

Proposition 2.1. *For any positive matrix $K(p,q)$ on E, there exists a uniquely determined functional Hilbert space H_K comprising of functions $\{f\}$ on E and admitting the reproducing kernel $K(p,q)$ $(RKHS\ H_K)$ satisfying and characterized by*

$$K(\cdot,q) \in H_K \text{ for any } q \in E \tag{2.4}$$

and, for any $q \in E$ and for any $f \in H_K$

$$f(q) = (f(\cdot), K(\cdot,q))_{H_K}. \tag{2.5}$$

Then in particular, we have the following fundamental results:

(I) For the RKHS H_K admitting the reproducing kernel $K(p,q)$ defined by (2.3), the images $\{f(p)\}$ by (2.2) for $\mathcal{H}$ are characterized as the members of the RKHS H_K.

(II) In general, we have the inequality in (2.2)

$$\|f\|_{H_K} \leq \|\mathbf{f}\|_{\mathcal{H}}, \tag{2.6}$$

however, for any $f \in H_K$ there exists a uniquely determined $\mathbf{f}^* \in \mathcal{H}$ satisfying

$$f(p) = (\mathbf{f}^*, \mathbf{h}(p))_{\mathcal{H}} \quad \text{on} \quad E \tag{2.7}$$

and

$$\|f\|_{H_K} = \|\mathbf{f}^*\|_{\mathcal{H}}. \tag{2.8}$$

In (2.6), the isometry holds if and only if $\{\mathbf{h}(p); p \in E\}$ is complete in $\mathcal{H}$.

Here, we shall assume that

$$\{\mathbf{h}(p); p \in E\}$$

is complete in $\mathcal{H}$.

Therefore, in (2.2) we have the isometric identity

$$\|f\|_{H_K} = \|\mathbf{f}\|_{\mathcal{H}}. \tag{2.9}$$

Now, for any $\mathbf{f} \in \mathcal{H}$ we consider the approximation

$$\inf_{\|\mathbf{g}\|_{\mathcal{H}} \leq R} \|\mathbf{f} - \mathbf{g}\|. \tag{2.10}$$

Of course, if $\|\mathbf{f}\|_{\mathcal{H}} \leq R$, there is no problem, since (2.10) is zero. The best approximation $\mathbf{g}^*$ in (2.10) is given by

$$\mathbf{g}^* = \frac{R\mathbf{f}}{\|\mathbf{f}\|_{\mathcal{H}}}$$

and we obtain the result, as we see from Schwarz's inequality

$$\inf_{\|\mathbf{g}\|_{\mathcal{H}} \leq R} \|\mathbf{f} - \mathbf{g}\| = \|\mathbf{f}\|_{\mathcal{H}} \left(1 - \frac{R}{\|\mathbf{f}\|_{\mathcal{H}}}\right). \tag{2.11}$$

Now, we shall transform the estimate onto H_K by using the linear mapping (2.2) and the isometry (2.9) in the form

$$\inf_{\|\mathbf{g}\|_{\mathcal{H}} \leq R} \|f - g\|_{H_K} = \|\mathbf{f}\|_{\mathcal{H}} \left(1 - \frac{R}{\|\mathbf{f}\|_{\mathcal{H}}}\right) \tag{2.12}$$

for $\|\mathbf{f}\|_{\mathcal{H}} \geq R$.

In statistical learning theory, we have the linear mappings

$$f(p) = \int_T F(t)\overline{h(t,p)}dm(t), \tag{2.13}$$

where $F \in L_2(T, dm)$, and the integral kernels $h(t,p)$ on $T \times T$ are Hilbert-Schmidt kernels satisfying

$$\int\int_{T\times T} |h(t,p)|^2 dm(t)dm(p) < \infty. \tag{2.14}$$

In our formulation, this will mean that the images of (2.2) for $\mathcal{H}$ belong to $\mathcal{H}$ again. Then, in the estimate (2.12) we will be able to consider a more concrete norm $\mathcal{H}$ than H_K. However, this part will be very delicate in the exact estimate of the norms in $\mathcal{H}$ and H_K, as we can see from the following examples.

Indeed, for our integral transforms in Theorems 1.1 and 1.2, the associated reproducing kernel Hilbert spaces are realized as follows:

In Theorem 1.1,

$$\|f\|_{H_K}^2 = \sum_{j=0}^{\infty} \frac{1}{j!\Gamma(j+2q+1)} \int_0^\infty |\partial_x^j (x(f'(x))|^2 x^{2j+2q-1} dx. \tag{2.15}$$

In Theorem 1.2,

$$\|u_F(\cdot,t)\|_{H_K}^2 = \sum_{j=0}^{\infty} \frac{(2t)^j}{j!} \int_{-\infty}^{\infty} |\partial_x^j u_F(x,t)|^2 dx. \tag{2.16}$$

In Theorem 1.3, we obtain the inequality

$$\int_{-\infty}^{\infty} |f(x)|^2 \exp(-a^2x^2)dx \le \frac{\sqrt{2\pi}}{a}\|f\|_{\mathbf{F}_a}^2. \tag{2.17}$$

See [2] for these results. In particular, note that inequality (2.17) is not trivial at all.

We take the simplest parts $j = 0$ in those realizations of the associated reproducing kernel Hilbert spaces. Then, we obtain Theorems 1.1 and 1.2. From (2.17), we have Theorem 1.3. This part of the theorems depends on case by case arguments and so, the results obtained are intricate.

3. INCOMPLETE CASE

In the case that $\{\mathbf{h}(p); p \in E\}$ is not complete in $\mathcal{H}$, similar estimates generally hold. We shall give a typical example.

We shall consider the integral transform, for any fixed x

$$u_F(x,t) = \frac{1}{2\pi}\int_{x-ct}^{x+ct} F(\xi)d\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi)\theta(ct - |x-\xi|)d\xi, \tag{3.1}$$

which gives the solution of the wave equation

$$\frac{\partial^2 u_F(x,t)}{\partial t^2} = c^2 \frac{\partial^2 u_F(x,t)}{\partial x^2} \quad (c > 0, \quad \text{constant})$$

satisfying

$$\frac{\partial u_F(x,t)}{\partial t}|_{t=0} = F(x), u(x,0) = 0 \quad \text{on} \quad \mathbf{R}.$$

Then we have the identity

$$2c\int_0^\infty \left(\frac{\partial u_F(x,t)}{\partial t}\right)^2 dt = \min \int_{-\infty}^{\infty} F(\xi)^2 d\xi \tag{3.2}$$

where the minimum is taken over all the functions $F \in L_2(\mathbf{R}) = L_2$ satisfying

$$u_F(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi)\theta(ct - |x-\xi|)d\xi,$$

for all $t > 0$ ([2, pp. 143-147]). Then, we obtain

Theorem 3.1. *In the integral transform (3.1) for L_2 functions F, we have the inequality*

$$\inf_{\|G\|_{L_2}\le R} \sqrt{2c}\|\frac{\partial u_F(x,t)}{\partial t} - \frac{\partial u_G(x,t)}{\partial t}\|_{L_2(0,\infty)} \ge \|F\|_{L_2}\left(1 - \frac{R}{\|F\|_{L_2}}\right)$$

for $\|F\|_{L_2} \le R$.

REFERENCES

[1] F. CUCKER AND S. SMALE, On the mathematical foundations of learning, *Bull. of Amer. Math. Soc.*, **39** (2001), 1–49.

[2] S. SAITOH, *Integral Transforms, Reproducing Kernels and Their Applications*, Pitman Res. Notes in Math. Series, **369** (1997), Addison Wesley Longman, UK.

[3] S. SAITOH, New type approximation error and convergence rate estimates in statistical learning theory, *J. Analysis and Applications*, **1** (2003), 33–39.

[4] S. SMALE AND D. X. ZHOU, Estimating the approximation error in learning theory, *Anal. Appl.*, **1** (2003), 17–41.