The Complexity of Automatic Algorithms Employing Continuous Linear Functionals

Fred J. Hickernell

Room E1-208, Department of Applied Mathematics, Illinois Institute of Technology, 10 W. 32nd St., Chicago, IL 60616

Abstract

Keywords:

1. The Basic Problem

Let $\mathcal{H}_{\text{in}}$ be a separable Banach space of input functions with basis $\{u_i\}_{i\in\mathcal{I}}$, where $\mathcal{I}$ is a countable index set, and let the norm for this Banach space be defined as the $\ell_{p_{\text{in}}}$-norm terms of the Fourier coefficients as follows:

$$f = \sum_{i\in\mathcal{I}} \hat{f}_i u_i \in \mathcal{H}_{\text{in}}, \qquad \|f\|_{\mathcal{H}_{\text{in}}} = \left\|\left(\hat{f}_i\right)_{i\in\mathcal{I}}\right\|_{p_{\text{in}}}, \qquad 1 \le p_{\text{in}} \le \infty.$$

Here the $\hat{f}_i$ denotes the i^{th} Fourier coefficient of a function $f \in \mathcal{H}_{\text{in}}$. If $p_{\text{in}} = 2$, then $\mathcal{H}_{\text{in}}$ is a Hilbert space. Similarly, let $\mathcal{H}_{\text{out}}$ be a separable Banach space of outputs with basis $\{v_i\}_{i\in\mathcal{I}}$, whose norm may be defined as follows:

$$f = \sum_{i\in\mathcal{I}} \tilde{f}_i v_i \in \mathcal{H}_{\text{out}}, \qquad \|f\|_{\mathcal{H}_{\text{in}}} = \left\|\left(\tilde{f}_i\right)_{i\in\mathcal{I}}\right\|_{p_{\text{out}}}, \qquad 1 \le p_{\text{out}} \le \infty.$$

In this case $\tilde{f}_i$ denotes the i^{th} Fourier coefficient of a function $f \in \mathcal{H}_{\text{out}}$. Define the solution operator S by $S(u_i) = \lambda_i v_i$, and so

$$S(f) = \sum_{i\in\mathcal{I}}^{\infty} \lambda_i \hat{f}_i v_i, \qquad \forall f \in \mathcal{H}_{\text{in}}. \tag{1}$$

Here the λ_i are the bounded singular values of the operator S. A generalization of Hölder's inequality is

$$\|(a_i b_i)_{i\in\mathcal{I}}\|_r \le \|(a_i)_{i\in\mathcal{I}}\|_p \, \|(b_i)_{i\in\mathcal{I}}\|_q \,, \qquad 1 \le r, p \le \infty, \quad q = \frac{pr}{\max(p-r, 0)}. \tag{2}$$

Thus, to ensure that the solution operator is bounded, it is assumed that

$$\|(\lambda_i)_{i\in\mathcal{I}}\|_q < \infty, \quad q = \frac{p_{\text{in}}p_{\text{out}}}{\max(p_{\text{in}} - p_{\text{out}}, 0)}. \tag{3}$$

This problem definition is rather general in some ways, e.g., allowing the inputs and outputs to lie in Banach spaces, rather than Hilbert spaces. However, a key requirement is that $S(u_i) = \lambda_i v_i$ for the solution operator of interest, which is a serious restriction on the choice of bases. Moreover, the norms of the Banach spaces cannot be arbitrarily defined, but must be defined in terms of the coefficients of the series expansions of the inputs and outputs.

For example, if $\mathcal{H}_{\text{in}} = \mathcal{H}_{\text{out}} = \mathcal{L}_2[0,1]$, and $S : f \mapsto f$ is the embedding operator, then one may choose a trigonometric polynomial basis:

$$\|f\|_{\mathcal{H}_{\text{in}}} = \|f\|_{\mathcal{H}_{\text{in}}} = \left[\int_0^1 |f(x)|^2 \, \mathrm{d}x\right]^{1/2},$$
$$u_i(x) = v_i(x) = \mathrm{e}^{2\pi\sqrt{-1}ix}, \quad \lambda_i = 1, \qquad i \in \mathbb{Z},$$
$$p_{\text{in}} = p_{\text{out}} = 2, \quad q = \infty.$$

If S computes the derivative of a function, $S : f \mapsto f'$, then the space $\mathcal{H}_{\text{in}}$ must have a stronger inner product to ensure that the λ_i are bounded, e.g.,

$$\|f\|_{\mathcal{H}_{\text{in}}} = \left[\int_0^1 \left\{|f(x)|^2 + |f'(x)|^2\right\} \mathrm{d}x\right]^{1/2},$$
$$u_i(x) = \frac{\mathrm{e}^{2\pi\sqrt{-1}ix}}{\sqrt{1+4\pi i^2}}, \quad \lambda_i = \frac{2\pi\sqrt{-1}i}{\sqrt{1+4\pi i^2}}, \qquad i \in \mathbb{Z},$$
$$p_{\text{in}} = p_{\text{out}} = 2, \quad q = \infty.$$

Suppose that one may choose arbitrary linear functionals to obtain data, and let $i_1, i_2, \ldots$ be an ordering of the elements in $\mathcal{I}$. Then one may approximate $S(f)$ by the first n terms of its infinite series representation:

$$A_n(f) = \sum_{j=1}^{n} \lambda_{i_j} \hat{f}_{i_j} v_{i_j} \qquad \forall f \in \mathcal{H}_{\text{in}},\ n \in \mathbb{N}, \tag{4a}$$

$$\|S(f) - A_n(f)\|_{\mathcal{H}_{\text{out}}} = \left\|\left(\lambda_{i_j}\hat{f}_{i_j}\right)_{j\ge n+1}\right\|_{p_{\text{out}}} \tag{4b}$$

Again by the generalization of Hölder's inequality, (2), it follows that

$$\sup_{0\ne f\in\mathcal{H}_{\text{in}}} \frac{\|S(f) - A_n(f)\|_{\mathcal{H}_{\text{out}}}}{\|f\|_{\mathcal{H}_{\text{in}}}} = \left\|\left(\lambda_{i_j}\right)_{j\ge n+1}\right\|_q, \tag{4c}$$

where q was defined in (3). For the example above, a natural ordering would be $i_j = (-1)^{j-1}\lceil (j-1)/2 \rceil$. An automatic algorithm for approximating $S(f)$ must have a way of reliably bounding

$$\left\|\left(\lambda_{i_j}\hat{f}_{i_j}\right)_{j\ge n+1}\right\|_{p_{\text{out}}} \quad \text{or} \quad \left\|\left(\lambda_{i_j}\right)_{j\ge n+1}\right\|_\infty$$

in terms of function data. The next two sections describe two possible ways to do this.

2. Approximating a Weaker Norm as a Surrogate for a Stronger Norm

The first method for constructing an automatic algorithm assumes two Banach subspaces of input functions, $\mathcal{F} \subseteq \mathcal{G} \subseteq \mathcal{H}_{\text{in}}$, whose norms are defined as weighted $\ell_{p_{\text{in}}}$-norms of the series coefficients. Specifically, let $(\nu_i)_{i\in\mathcal{I}}$ and $(\omega_i)_{i\in\mathcal{I}}$ be positive sequences of weights, and

$$f = \sum_{i\in\mathcal{I}} \hat{f}_i u_i \in \mathcal{H}_{\text{in}}, \qquad \|f\|_{\mathcal{G}} = \left\| \left(\frac{\hat{f}_i}{\omega_i} \right)_{i\in\mathcal{I}} \right\|_{p_{\text{in}}}, \qquad \|f\|_{\mathcal{F}} = \left\| \left(\frac{\hat{f}_i}{\nu_i \omega_i} \right)_{i\in\mathcal{I}} \right\|_{p_{\text{in}}}.$$

If $\inf_{i\in\mathcal{I}} \omega_i = 0$, then the $\mathcal{G}$-norm is stronger than the $\mathcal{H}_{\text{in}}$-norm. If $\inf_{i\in\mathcal{I}} \nu_i = 0$, then the $\mathcal{F}$-norm is stronger than the $\mathcal{G}$-norm. For these two new norms the worst-case error can be bounded tightly as in (4c) as follows:

$$\begin{aligned} \sup_{0\neq f\in\mathcal{G}} \frac{\|S(f) - A_n(f)\|_{\mathcal{H}_{\text{out}}}}{\|f\|_{\mathcal{G}}} &= \left\| \left(\omega_{i_j} \lambda_{i_j} \right)_{j\geq n+1} \right\|_q, \\ \sup_{0\neq f\in\mathcal{F}} \frac{\|S(f) - A_n(f)\|_{\mathcal{H}_{\text{out}}}}{\|f\|_{\mathcal{F}}} &= \left\| \left(\nu_{i_j} \omega_{i_j} \lambda_{i_j} \right)_{j\geq n+1} \right\|_q, \end{aligned} \tag{5}$$

where again q is given in terms of p_{in} and p_{out} by (3).

The (weaker) $\mathcal{G}$-semi-norm of a function $f \in \mathcal{F}$ may be estimated using the function data $\langle u_{i_j}, f \rangle_{\mathcal{H}_{\text{in}}}$ that is also used to approximate the solution $S(f)$ as follows:

$$G_n(f) = \left\| \sum_{i=1}^{n} \hat{f}_{i_j} u_{i_j} \right\|_{\mathcal{G}} = \left\| \left(\frac{\hat{f}_{i_j}}{\omega_{i_j}} \right)_{j=1}^{n} \right\|_{p_{\text{in}}}. \tag{6}$$

The error of this approximation is given by the two-sided inequality

$$\begin{aligned} 0 \leq \|f\|_{\mathcal{G}}^{p_{\text{in}}} - G_n^{p_{\text{in}}}(f) &= \left\| \left(\frac{\hat{f}_{i_j}}{\omega_{i_j}} \right)_{j\geq n+1} \right\|_{p_{\text{in}}}^{p_{\text{in}}} \\ &\leq \left\| \left(\nu_{i_j} \right)_{j\geq n+1} \right\|_{\infty}^{p_{\text{in}}} \|f\|_{\mathcal{F}}^{p_{\text{in}}}, \quad 1 \leq p_{\text{in}} < \infty. \end{aligned}$$

Assuming f lies in the specified cone of functions, defined by

$$\mathcal{C}_\tau = \{ f \in \mathcal{F} : \|f\|_{\mathcal{F}} \leq \tau \|f\|_{\mathcal{G}} \}, \tag{7}$$

the arguments of [1] lead to a two-sided bound on the $\mathcal{G}$-semi-norm f in terms of $G_n(f)$. Specifically, if $\left\| \left(\nu_{i_j} \right)_{j\geq n+1} \right\|_{\infty} < 1/\tau$, then for all $f \in \mathcal{C}_\tau$, it follows that

$$G_n(f) \leq \|f\|_{\mathcal{G}} \leq \mathfrak{C} G_n(f)$$

$$\mathfrak{C} = \begin{cases} \left[1 - \tau^{p_{\text{in}}} \left\| \left(\nu_{i_j} \right)_{j\geq n+1} \right\|_{\infty}^{p_{\text{in}}} \right]^{-1/p_{\text{in}}}, & 1 \leq p_{\text{in}} < \infty, \\ 1, & p_{\text{in}} = \infty. \end{cases}$$

The case $p_{\text{in}} = \infty$ follows from the $p_{\text{in}} < \infty$ case by taking the limit as $p_{\text{in}} \to \infty$. Alternatively, the case $p_{\text{in}} = \infty$ may be proven directly by choosing a positive δ strictly less than $1/\big(\tau \big\| \left(\nu_{i_j}\right)_{j \ge n+1} \big\|_{\infty}\big)$ and then noting that there exists $j^* \ge n+1$ with

$$\left\| \left(\frac{\hat{f}_{i_j}}{\omega_{i_j}} \right)_{j \ge n+1} \right\|_{\infty} \le \left| \frac{\hat{f}_{i_{j^*}}}{\omega_{i_{j^*}}} \right| (1+\delta)$$

$$< \frac{1}{\tau \big\| \left(\nu_{i_j}\right)_{j \ge n+1} \big\|_{\infty}} \left| \frac{\hat{f}_{i_{j^*}}}{\omega_{i_{j^*}}} \right| \le \frac{1}{\tau} \left| \frac{\hat{f}_{i_{j^*}}}{\nu_{i_{j^*}} \omega_{i_{j^*}}} \right|$$

$$\le \frac{1}{\tau} \left\| \left(\frac{\hat{f}_{i_j}}{\nu_{i_j} \omega_{i_j}} \right)_{j \ge n+1} \right\|_{\infty} \le \frac{\|f\|_{\mathcal{F}}}{\tau} \le \|f\|_{\mathcal{G}}$$

This strict inequality implies the existence of some $j \le n$ such that $|\hat{f}_{i_j}/\omega_{i_j}| = \|f\|_{\mathcal{G}}$, and so $G_n(f)$ is has no error.

Combining this upper bound on $\|f\|_{\mathcal{G}}$, the cone condition, and error bound (5), implies that

$$\|S(f) - A_n(f)\|_{\mathcal{H}_{\text{out}}} \le \tau \mathfrak{C} G_n(f) \left\| \left(\nu_{i_j} \omega_{i_j} \lambda_{i_j}\right)_{j \ge n+1} \right\|_{\infty}.$$

Thus, one increases n until the right hand side falls below some tolerance, ε. The drawback of this approach is that one needs to choose the weights ν_i and ω_i, which define the spaces $\mathcal{F}$ and $\mathcal{G}$.

3. Assuming a Gentle Decay in the Terms of the Series

Another method for estimating the error of A_n assumes that the decay of the Fourier coefficients of f follows some general rate of decay, which need not be known precisely. Suppose again that we have some ordering of the linear functionals, as in the previous section, and let $0 = n_0 < n_1 < n_2 < \ldots$ be an ordered, unbounded sequence of integers. Define the sums

$$\sigma_k(f) = \left\| \left(\lambda_{i_j} \hat{f}_{i_j} \right)_{j=n_{k-1}+1}^{n_k} \right\|_{p_{\text{out}}}, \qquad k = 1, 2, \ldots \tag{8}$$

Given two numbers greater than one, s_1 and s_2, define the cone of functions in $\mathcal{H}_{\text{in}}$ as those whose Fourier coefficients decay at a given rate:

$$\mathcal{C} = \left\{ f \in \mathcal{H}_{\text{in}} : \sigma_k(f) \le s_1 s_2^{\kappa - k} \sigma_\kappa(f), \ \forall \kappa, k \in \mathbb{N}, \ \kappa \le k \right\}. \tag{9}$$

To see how these conditions might arise naturally, consider the case where the terms in the sum defining $\sigma_k(f)$ decay *algebriacally* and the n_k increase *geometrically*:

$$C_{\text{lo}} j^{-p} \le \left| \lambda_{i_j} \left\langle u_{i_j}, f \right\rangle_{\mathcal{H}_{\text{in}}} \right| \le C_{\text{up}} j^{-p}, \quad p > 1, \ j \in \mathbb{N},$$
$$n_k = a b^k \quad a, b, k \in \mathbb{N}, \ b \ge 2.$$

The sum of positive integers raised to a power can be interpreted as a left or right rectangle rule for approximating an integral. This leads to upper and lower bounds for the sums:

$$\sum_{j=n_{\mathrm{lo}}}^{n_{\mathrm{up}}} j^{-p} \geq \int_{n_{\mathrm{lo}}}^{n_{\mathrm{up}}+1} x^{-p}\,\mathrm{d}x = \frac{n_{\mathrm{lo}}^{1-p} - (n_{\mathrm{up}}+1)^{1-p}}{p-1},$$
$$\sum_{j=n_{\mathrm{lo}}}^{n_{\mathrm{up}}} j^{-p} \leq \int_{n_{\mathrm{lo}}-1}^{n_{\mathrm{up}}} x^{-p}\,\mathrm{d}x = \frac{(n_{\mathrm{lo}}-1)^{1-p} - n_{\mathrm{up}}^{1-p}}{p-1}.$$

These two bounds can be used to prove that f lies in the cone (9) for appropriately chosen s_1 and s_2:

$$\sigma_k(f) \geq C_{\mathrm{lo}} \left\{ \frac{(n_{k-1}+1)^{1-pp_{\mathrm{out}}} - (n_k+1)^{1-pp_{\mathrm{out}}}}{pp_{\mathrm{out}}-1} \right\}^{1/p_{\mathrm{out}}}$$
$$= C_{\mathrm{lo}} \left\{ \frac{[ab^{k-1}]^{1-pp_{\mathrm{out}}}}{pp_{\mathrm{out}}-1} \left[(1+a^{-1}b^{1-k})^{1-pp_{\mathrm{out}}} - (b+a^{-1}b^{1-k})^{1-pp_{\mathrm{out}}} \right] \right\}^{1/p_{\mathrm{out}}},$$

$$\sigma_k(f) \leq C_{\mathrm{up}} \left\{ \frac{n_{k-1}^{1-pp_{\mathrm{out}}} - n_k^{1-pp_{\mathrm{out}}}}{pp_{\mathrm{out}}-1} \right\}^{1/p_{\mathrm{out}}}$$
$$= C_{\mathrm{up}} \left\{ \frac{[ab^{k-1}]^{1-pp_{\mathrm{out}}}(1-b^{1-pp_{\mathrm{out}}})}{pp_{\mathrm{out}}-1} \right\}^{1/p},$$

$$\frac{\sigma_k(f)}{\sigma_\kappa(f)} \leq \frac{C_{\mathrm{up}}}{C_{\mathrm{lo}}} \left\{ \frac{(1-b^{1-pp_{\mathrm{out}}})b^{(pp_{\mathrm{out}}-1)(\kappa-k)}}{(1+a^{-1}b^{1-\kappa})^{1-pp_{\mathrm{out}}} - (b+a^{-1}b^{1-\kappa})^{1-pp_{\mathrm{out}}}} \right\}^{1/p_{\mathrm{out}}}$$
$$\leq \frac{C_{\mathrm{up}}}{C_{\mathrm{lo}}} \left\{ \frac{(1-b^{1-pp_{\mathrm{out}}})b^{(pp_{\mathrm{out}}-1)(\kappa-k)}}{(1+a^{-1})^{1-pp_{\mathrm{out}}} - (b+a^{-1})^{1-pp_{\mathrm{out}}}} \right\}^{1/p_{\mathrm{out}}} = s_1 s_2^{\kappa-k},$$

where

$$s_1 = \frac{C_{\mathrm{up}}}{C_{\mathrm{lo}}} \left\{ \frac{(1-b^{1-pp_{\mathrm{out}}})}{(1+a^{-1})^{1-pp_{\mathrm{out}}} - (b+a^{-1})^{1-pp_{\mathrm{out}}}} \right\}^{1/p_{\mathrm{out}}}, \qquad s_2 = b^{p-1/p_{\mathrm{out}}}.$$

One may also consider the case where the terms in the sum defining $\sigma_k(f)$ decay *exponentially* and the n_k increase *linearly*:

$$C_{\mathrm{lo}} p^{-j} \leq \left| \lambda_{i_j} \left\langle u_{i_j}, f \right\rangle_{\mathcal{H}_{\mathrm{in}}} \right| \leq C_{\mathrm{up}} p^{-j}, \quad p > 1,\ j \in \mathbb{N},$$
$$n_k = a + kb \quad a, b, k \in \mathbb{N}.$$

The geometric sum that now arises in a bound on the definition of $\sigma_k(f)$ takes the form

$$\sum_{j=n_{\mathrm{lo}}}^{n_{\mathrm{up}}} p^{-j} = \frac{p^{-n_{\mathrm{lo}}} - p^{-n_{\mathrm{up}}-1}}{1-p^{-1}},$$

which implies that

$$\begin{aligned}
\frac{\sigma_k(f)}{\sigma_\kappa(f)} &\le \frac{C_{\text{up}}}{C_{\text{lo}}} \left\{ \frac{p^{-p_{\text{out}}(n_{k-1}-1} - p^{-p_{\text{out}}(n_k-1)}]}{C_{\text{lo}}^2 [p^{-p_{\text{out}}(n_{\kappa-1}-1)} - p^{-p_{\text{out}}(n_\kappa-1)}]} \right\}^{1/p_{\text{out}}} \\
&= \frac{C_{\text{up}} p^{n_{\kappa-1}-n_{k-1}}}{C_{\text{lo}}} \left\{ \frac{1 - p^{p_{\text{out}}(n_{k-1}-n_k)}}{1 - p^{p_{\text{out}}(n_{\kappa-1}-n_\kappa)}]} \right\}^{1/p} \\
&= \frac{C_{\text{up}} p^{b(\kappa-k)}}{C_{\text{lo}}} = s_1 s^{\kappa-k},
\end{aligned}$$

where

$$s_1 = \frac{C_{\text{up}}}{C_{\text{lo}}}, \qquad s_2 = p^b.$$

From the definition of the cone and (4b), one can show that

$$\begin{aligned}
\|S(f) - A_{n_\kappa}(f)\|_{\mathcal{H}_{\text{out}}} &= \left\| \left(\lambda_{i_j} \hat{f}_{i_j} \right)_{j \ge n_\kappa+1} \right\|_{p_{\text{out}}} \\
&= \left\{ \sum_{k=\kappa+1}^{\infty} \sum_{j=n_{k-1}+1}^{n_k} \left| \lambda_{i_j} \tilde{f}_{i_j} \right|^{p_{\text{out}}} \right\}^{1/p_{\text{out}}} \\
&= \left\{ \sum_{k=\kappa+1}^{\infty} \sigma_k^{p_{\text{out}}}(f) \right\}^{1/p_{\text{out}}} \\
&\le \left\{ \sum_{k=\kappa+1}^{\infty} [\sigma_\kappa(f) s_1 s_2^{\kappa-k}]^{p_{\text{out}}} \right\}^{1/p_{\text{out}}} \\
&= \frac{\sigma_\kappa(f) s_1 s_2}{[1 - s_2^{-p_{\text{out}}}]^{1/p_{\text{out}}}}
\end{aligned}$$

Since the right hand side depends on the data, we have a data-driven error bound. Note also that $\sigma_\kappa(f)$ decays as quickly with respect to n as the true error. Unlike the earlier method, one does not need to know the decay rate.

[1] Y. Ding, N. Clancy, C. Hamilton, F. J. Hickernell, Y. Zhang, The complexity of guaranteed automatic algorithms, in preparation (2013).