Proceedings of the World Congress on Engineering 2008 Vol II
WCE 2008, July 2 - 4, 2008, London, U.K.

Linear Combinations of Gaussians with a Single Variance are Dense in L^2

Craig Calcaterra[1]

Abstract—**Linear combinations of translations of e^{-x^2} are shown to be dense in $L^p(\mathbb{R})$ for $1 \leq p < \infty$ with a nearly trivial proof. A potential application to signal analysis is detailed where the Gaussian filter is seen to be a universal synthesizer with arbitrarily short load time.**

Key Words: Gaussian filter, Hermite approximation

The Result

Denote the space of square integrable functions $f : \mathbb{R} \to \mathbb{R}$ as $L^2(\mathbb{R})$ with norm $\|f\|_2 := \sqrt{\int_{\mathbb{R}} |f(x)|^2 \, dx}$. We use $f \underset{\epsilon}{\approx} g$ to mean $\|f - g\|_2 < \epsilon$.

Theorem 1 *For any $f \in L^2(\mathbb{R})$ and any $\epsilon > 0$ there exists $t > 0$ and $N \in \mathbb{N} = \{0, 1, 2, ...\}$ and $a_n \in \mathbb{R}$ such that*

$$f \underset{\epsilon}{\approx} \sum_{n=0}^{N} a_n e^{-(x-nt)^2}.$$

Proof. Since the span of the Hermite functions is dense in $L^2(\mathbb{R})$ we have for some N

$$f \underset{\epsilon/2}{\approx} \sum_{n=0}^{N} b_n \frac{d^n}{dx^n}\left(e^{-x^2}\right).$$

Now use finite backward differences. We have for some small $t > 0$

$$\sum_{n=0}^{N} b_n \frac{d^n}{dx^n}\left(e^{-x^2}\right) \underset{\epsilon/2}{\approx} \sum_{n=0}^{N} b_n \frac{1}{t^n} \sum_{k=0}^{n} (-1)^k \binom{n}{k} e^{-(x-kt)^2}.$$

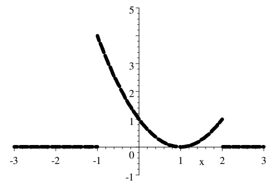
■

So all signals can be approximated by **Gaussian bumps**. This bump approximation scheme is an affront to our intuition; it promises we can approximate to any degree of accuracy a function such as the following characteristic function of an interval

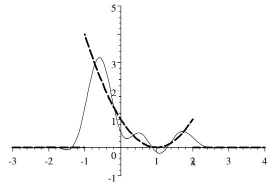
$$\chi_{[-101,-100]}(x) := \begin{cases} 1 & \text{for } x \in [-101, -100] \\ 0 & \text{otherwise} \end{cases}$$

with support far from the means of the Gaussians $e^{-(x-nt)^2}$ which shrink precipitously away from $nt \in [0, \infty)$. Theorem 1 also promises an approximation of functions with any small variance such as e^{-100x^2} despite the fact that $e^{-(x+nt)^2}$ all have an identical variance.

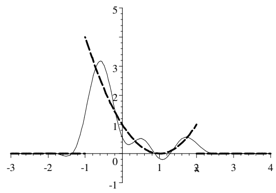
The following figures give graphical evidence for Theorem 1. The original function is the dashed graph and the approximation is a solid curve.



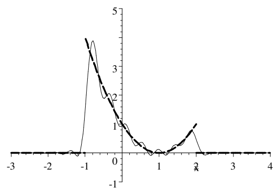
$f(x) = (x-1)^2 \chi_{[-1,2]}(x)$



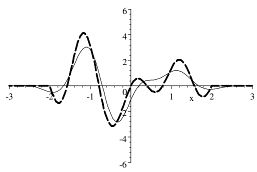
Hermite series $N = 20$



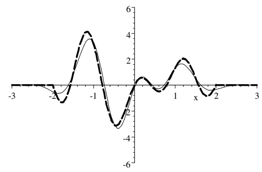
Bump approximation
$N = 20$, $t = .01$



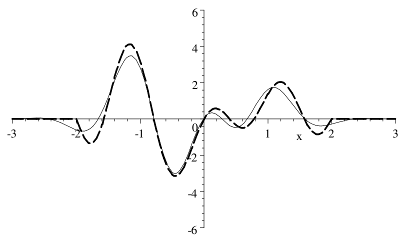
Bump approximation
$N = 100$, $t = .001$



Hermite $N = 12$



Hermite $N = 15$



Bump approximation
$N = 15$, $t = .01$

Clearly Hermite expansions perform better with smooth data; see [4], [8], e.g. The first function approximated above is discontinuous at two points, and we need large values of N for visual accuracy. The second is still non-smooth at two points, but needs much less computation.

[1] Submitted: March 4, 2008; Department of Mathematics, Metropolitan State University, Saint Paul, MN 55106, USA Tel: (651) 793-1423 Email: craig.calcaterra@metrostate.edu Thanks to Axel Boldt and David Bleecker.

Application to Signal Analysis

The basic insight that propels signal analysis is that most any signal can be approximated by a linear combination of sine waves with countably infinitely many frequencies. Bump approximation gives an alternative signal approximation scheme to explore, using linear combinations of Gaussians with a single variance. We use a linear time-invariant (LTI) system to implement Theorem 1. An analog reading of the following results give an arbitrarily short load time; a digital implementation is naturally motivated which gives a simple compression scheme. The LTI system we need is the Gaussian filter which is represented with the operator $\mathcal{W} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by convolution with a Gaussian $G(x) := \frac{1}{\sqrt{4\pi}} e^{-x^2}$, that is,

$$\mathcal{W}(f)(x) := (f * G)(x) = \frac{1}{\sqrt{4\pi}} \int_{\mathbb{R}} f(s)\, e^{-(s-x)^2} ds.$$

The letter $\mathcal{W}$ is used to denote the operator because this is precisely the Weierstrass transform from pure mathematics. $\mathcal{W}$ is a low-pass frequency filter and in the 2-D case is famous as Gaussian blur in image processing. The question we wish to answer in this section is: "Given an arbitrary signal f what input for the Gaussian filter has output f?"

The simplest application of Theorem 1 is to use the Gaussian filter's impulse response. An impulse response for a system such as $\mathcal{W}$ is the output after feeding the system a Dirac delta distribution, δ_0. In fact since δ_0 is the identity under the convolution operation the impulse response is $\mathcal{W}(\delta_0) = \delta_0 * G = G$. The Gaussian filter takes its name from its impulse response.

$\mathcal{W}$ is an LTI system because it is linear

$$\mathcal{W}(af + g) = a\mathcal{W}(f) + \mathcal{W}(g)$$

and time-invariant

$$\mathcal{W}(f) \circ \tau_t = \mathcal{W}(f \circ \tau_t)$$

where $\tau_t(x) = x + t$. This gives us $\mathcal{W}(\delta_t)(x) = G(x - t)$ and Theorem 1 gives

Corollary 2 *For any $f \in L^2(\mathbb{R})$ and any $\epsilon > 0$ there exists $t > 0$ and $N \in \{0, 1, 2, ...\}$ and $a_n \in \mathbb{R}$ such that*

$$f \underset{\epsilon}{\approx} \mathcal{W}\left(\sum_{n=0}^{N} a_n \delta_{nt}\right).$$

So any signal is the Gaussian blur of a sum of pulses. The a_n are easily calculated analytically from the proof of Theorem 1. There are well-known applications of Fourier analysis to speed up the digital processing of a convolution–the Fourier transform of the convolution of two functions is the product of the transforms. The Gaussian filter is particularly simple since the transform of the Gaussian is another Gaussian.

There are, however, inherent difficulties in realizing an analog Gaussian filter. The first attempt is recorded in [2]. Then [3] details the need for n elements to make an nth order approximation to a perfect analog Gaussian filter; various improvements have been made in the intervening years. Analytical investigations along the lines above, using [1], give criteria for determining the viability of any such approximate Gaussian.

References

[1] Craig Calcaterra, Foliating Metric Spaces, arxiv.org

[2] S. Darlington, Synthesis and Reactance of 4-poles, J. Math. & Phys., 18, pp. 257-353, 1939

[3] Milton Dishal, Gaussian-Response Filter Design, Electrical Communication, Volume 36, No. 1, 1959, p. 3-26.

[4] David Gottlieb and Steven Orszag, "Numerical Analysis of Spectral Methods," SIAM, 1977.

[5] Leslie Greengard and Xiaobai Sun, A New Version of the Fast Gauss Transform, Documenta Mathematica, Extra Volume ICM 1998, III, pp. 575-584.

[6] Greg Leibon, Daniel Rockmore & Gregory Chirikjian, A Fast Hermite Transform with Applications to Protein Structure Determination, Proceedings of the 2007 international Workshop on Symbolic-Numeric Computation (London, Ontario, Canada, July 25 - 27, 2007). SNC '07. ACM, New York, NY, 117-124.

[7] J. Madrenas, M. Verleysen, P. Thissen, and J. L. Voz, A CMOS Analog Circuit for Gaussian Functions, IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing, Vol. 43, No. 1, January 1996.

[8] Gabor Szegö, "Orthogonal Polynomials," American Mathematical Society, 3rd ed., 1967.