Symbolic and Numerical Scientific Computing 2008, Hagenberg, Austria, July 24–26

Convergence Theory for Geddes-Newton Series Expansions of Positive Definite Kernels

Keith O. Geddes
School of Computer Science
University of Waterloo, Ontario, Canada

Joint Work with Frederick W. Chapman (PhD, Waterloo, 2003)

Function Space Norms and Convergence

Assume $f : X \to \mathbb{C}$ and $f_n : X \to \mathbb{C}$ and $1 \leq p < \infty$.

- The **infinity norm of** f is $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

- The **Lebesgue** p**-norm of** f is $\|f\|_p = \left(\int_X |f(x)|^p \, dx \right)^{1/p}$.

- We say $f_n \to f$ **uniformly** if $\|f - f_n\|_\infty \to 0$ as $n \to \infty$.

- We say $f_n \to f$ **in** L^p if $\|f - f_n\|_p \to 0$ as $n \to \infty$.

Kernel Functions, Squares, and Diagonals

Let A be an *arbitrary set* and $f : A^2 \longrightarrow \mathbb{C}$ be an *arbitrary function*.

- We say that f **is a kernel on** A.

- Every kernel on A is defined on the **square** $A^2 = A \times A$.

- The **diagonal of the kernel** f is the *function*

$$\hat{f} : A \longrightarrow \mathbb{C} \text{ with } \hat{f}(a) = f(a,a) \text{ for all } a \in A.$$

- The **diagonal of the square** A^2 is the *subset*

$$\text{diag}(A^2) = \Big\{(a,a) \in A^2 : a \in A\Big\} \subset A^2.$$

Positive Definite Kernels

Let f be a kernel on A. (That is, assume $f : A^2 \to \mathbb{C}$.)

- For any $n \geq 1$ and $a_1, \ldots a_n \in A$, define the **Gram matrix of f at** $a_1, \ldots a_n$ by

$$G(f \mid a_1, \ldots a_n) = \Big[f(a_i, a_j)\Big]_{i,j=1}^{n} \in \mathbb{C}^{n \times n}.$$

- We call f a **positive definite kernel (PDK)** if $G(f \mid a_1, \ldots a_n)$ is a positive *semidefinite* matrix for all $n \geq 1$ and $a_1, \ldots a_n \in A$.

- We call f a **strictly positive definite kernel (SPDK)** if $G(f \mid a_1, \ldots a_n)$ is a positive *definite* matrix whenever $a_1, \ldots a_n \in A$ are *distinct*.

Two Ways to Construct Positive Definite Kernels

- **Positive Definite Functions:** If V is a vector space over $\mathbb{C}$, any function $\phi : V \to \mathbb{C}$ defines a translation invariant kernel f on V by

  $$f(x, y) = \phi(x - y).$$

  By definition, ϕ is a **(strictly) positive definite function on** V if f is a (strictly) positive definite kernel on V. Bochner's theorem characterizes all *continuous* positive definite functions ϕ via their Fourier transforms.

- **Inner Product Spaces:** Let $\langle \bullet, \bullet \rangle$ be an inner product on V. Any function $\Phi : A \to V$ defines a positive definite kernel f on A by

  $$f(x, y) = \langle \Phi(x), \Phi(y) \rangle .$$

  For example, $f(x, y) = \langle x, y \rangle$ defines a positive definite kernel on V (let $A = V$ and $\Phi = I$).

Examples of Strictly Positive Definite Kernels

- **"Hat Function" Kernel** (continuous but not differentiable on $\mathbb{R}^2$):

$$1 - |x - y| \text{ is a SPDK on } [0, 1]$$

- **Bessel Function Kernel** (an entire function on $\mathbb{C}^2$):

$$J_0(x - y) \text{ is a SPDK on } \mathbb{R}$$

- **Gaussian Kernel** (an entire function on $\mathbb{C}^2$):

$$\exp(-(x - y)^2) \text{ is a SPDK on } \mathbb{R}.$$

Why Are Positive Definite Kernels Important?

Positive definite kernels are of great theoretical interest, have many practical applications, and arise often in active areas of research, such as:

- kernel-based methods for machine learning
- covariance functions in mathematical statistics
- reproducing kernel Hilbert spaces in functional analysis
- radial basis functions for multivariate interpolation and approximation.

Properties of Positive Definite Kernels

Let f be a positive definite kernel on A.

- f is **Hermitian:** $f(a,b) = \overline{f(b,a)}$.
- f is *real and nonnegative over the diagonal:* $\hat{f}(a) = f(a,a) \geq 0$.
- f has the **Cauchy-Schwarz property:** $|f(a,b)| \leq \sqrt{\hat{f}(a)} \cdot \sqrt{\hat{f}(b)}$
- f has the **diagonal property:** $\|f\|_\infty = \left\|\hat{f}\right\|_\infty$
- f has the **integrability property:** $\|f\|_{2p} \leq \left\|\hat{f}\right\|_p$ for $1 \leq p < \infty$.

Applications to Geddes-Newton Series

Let f be a positive definite kernel on A. For all $n \geq 0$, let $r_n = f - s_n$, where s_n is the Geddes-Newton series expansion of f with n distinct diagonal splitting points $\{(a_i, a_i)\}_{i=0}^{n-1} \subset \text{diag}(A^2)$. We have the following three conclusions:

- r_n is a positive definite kernel on A for all $n \geq 0$. (Proof by induction on n: Use *Schur determinant formula* to derive key identity for Gram matrices

  $$\det G(r_{n+1} \mid b_1, \ldots, b_m) = \det G(r_n \mid a_n, b_1, \ldots, b_m)/r_n(a_n, a_n).$$

  Apply to *principal minor characterization* of positive semidefinite matrices.)

- $s_n \to f$ uniformly on A^2 if and only if $\hat{s}_n \to \hat{f}$ uniformly on A.

- If $\hat{s}_n \to \hat{f}$ in L^p on A, where $1 \leq p < \infty$, then $s_n \to f$ in L^{2p} on A^2.

Geddes-Newton Series Convergence Theorem

- **Theorem (Chapman & Geddes, 2008):** Assume $A \subset \mathbb{C}$ is compact and let f be a positive definite kernel on A. For all $n \geq 0$, let $r_n = f - s_n$, where s_n is the Geddes-Newton series expansion of f with n distinct diagonal splitting points $\{(a_i, a_i)\}_{i=0}^{n-1} \subset \text{diag}(A^2)$. If f is complex-analytic on a sufficiently large region containing A^2, then $s_n \to f$ absolutely and uniformly on A^2 at a linear rate or faster.

- **Proof Sketch:** Let $R_n = f - S_n$, where S_n is the *Boolean tensor product* which interpolates f on the grid lines $x = a_i$ and $y = a_i$ for $i = 0, \ldots n-1$ (see Cheney & Light, 2000). For all $n \geq 0$ and some fixed $\gamma \in (0, 1)$,

$$\|r_n\|_\infty = \|\hat{r}_n\|_\infty \leq \left\|\hat{R}_n\right\|_\infty \leq \|R_n\|_\infty = O(\gamma^n) \text{ as } n \to \infty.$$

  In addition, the Geddes-Newton series s_n converges absolutely by comparison with a geometric series with common ratio γ.

Convergence Theorems for Nonsmooth Kernels

- Parseval's identity is an absolutely convergent Geddes-Newton series expansion of the inner product on any separable Hilbert space.

- If f is a positive definite *continuous* kernel on a separable metric space A and the diagonal splitting points satisfy a *density hypothesis*, then the Geddes-Newton series expansion of f converges absolutely to f on A^2, with uniform convergence on every compact subset of A^2.

- If, in addition to the hypotheses above, the diagonal function $\hat{f} \in L^p(A)$, where $1 \leq p < \infty$, then the kernel $f \in L^{2p}(A^2)$ and the Geddes-Newton series expansion of f converges to f in L^{2p} on A^2. If A is a set of finite measure, then the Geddes-Newton series expansion of f also converges to f in L^1 on A^2.