Development of Kernels from Linear Functionals and Linear Operators

Stephen Vardeman
Analytics Iowa LLC
ISU Statistics and IMSE

A problem generalizing the smoothing spline problem

For constants $d_i > 0$, functionals F_i, and a linear differential operator L defined for continuous functions $w_k(x)$ by

$$L[h](x) = h^{(m)}(x) + \sum_{k=1}^{m-1} w_k(x) h^{(k)}(x)$$

Heckman considers the minimization of

$$\sum_{i=1}^{N} d_i (y_i - F_i[h])^2 + \lambda \int_a^b (L[h](x))^2 \, dx \tag{1}$$

in the space of functions

$$\mathcal{A} = \left\{ h : [a,b] \to \Re \middle| \begin{array}{c} \text{derivatives of } h \text{ up to order } m-1 \text{ are} \\ \text{absolutely continuous and } \int_a^b \left(h^{(m)}(x)\right)^2 dx < \infty \end{array} \right\}$$

An inner product for the function space

One may adopt an inner product for $\mathcal{A}$ of the form

$$\langle f, g \rangle_{\mathcal{A}} \equiv \sum_{k=0}^{m-1} f^{(k)}(a)\, g^{(k)}(a) + \int_a^b L[f](x)\, L[g](x)\, dx$$

and have a RKHS.

The assumption is made that the functionals F_i are continuous and linear, and thus that they are representable as $F_i[h] = \langle f_i, h \rangle_{\mathcal{A}}$ for some $f_i \in \mathcal{A}$. An important special case is that where $F_i[h] = h(x_i)$, but other linear functionals (e.g. $F_i[h] = \int_a^b H_i(x)\, h(x)\, dx$ for known H_i) have been used.

Development of the reproducing kernel

A reproducing kernel implied by this choice of inner product is derivable as follows.

First, there is a linearly independent set of functions $\{u_1, \ldots, u_m\}$ that is a basis for the subspace of $\mathcal{A}$ consisting of those elements h for which $L[h] = 0$ (the zero function). Call this subspace $\mathcal{A}_0$. The so-called Wronskian matrix associated with these functions is then

$$\underset{m\times m}{\mathbf{W}}(x) = \left(u_i^{(j-1)}(x)\right)$$

With

$$\mathbf{C} = \left(\mathbf{W}(a)\,\mathbf{W}(a)'\right)^{-1}$$

let

$$R_0(x,z) = \sum_{i,j} C_{ij}\, u_i(x)\, u_j(z)$$

Kernel development contd.

Further, there is a so-called Green's function associated with the operator L, a function $G(x, z)$ such that for all $h \in \mathcal{A}$ satisfying $h^{(k)}(a) = 0$ for $k = 0, 1, \ldots, m-1$

$$h(x) = \int_a^b G(x, z)\, L[h](z)\, dz$$

Let

$$R_1(x, z) = \int_a^b G(x, t)\, G(z, t)\, dt$$

The reproducing kernel associated with the inner product and L is then

$$R(x, z) = R_0(x, z) + R_1(x, z)$$

Decomposition of the function space

As it turns out, $\mathcal{A}_0$ is a RKHS with reproducing kernel R_0 under the inner product $\langle f, g \rangle_0 = \sum_{k=0}^{m-1} f^{(k)}(a)\, g^{(k)}(a)$. Further, the subspace of $\mathcal{A}$ consisting of those h with $h^{(k)}(a) = 0$ for $k = 0, 1, \ldots, m-1$ (call it $\mathcal{A}_1$) is a RKHS with reproducing kernel $R_1(x, z)$ under the inner product $\langle f, g \rangle_1 = \int_a^b L[f](x)\, L[g](x)\, dx$. Every element of $\mathcal{A}_0$ is $\perp$ to every element of $\mathcal{A}_1$ in $\mathcal{A}$ and every $h \in \mathcal{A}$ can be written uniquely as $h_0 + h_1$ for an $h_0 \in \mathcal{A}_0$ and an $h_1 \in \mathcal{A}_1$.

These facts can be used to show that a minimizer of (1) on Slide 2 exists and is of the form

$$h(x) = \sum_{k=1}^{m} \alpha_k u_k(x) + \sum_{i=1}^{N} \beta_i R_1(x_i, x)$$

Resulting matrix calculus problem

For $h(x)$ of this form, criterion (1) is of the form

$$(\mathbf{Y} - \mathbf{T}\boldsymbol{\alpha} - \mathbf{K}\boldsymbol{\beta})' \mathbf{D} (\mathbf{Y} - \mathbf{T}\boldsymbol{\alpha} - \mathbf{K}\boldsymbol{\beta}) + \lambda \boldsymbol{\beta}' \mathbf{K} \boldsymbol{\beta}$$

for

$$\mathbf{T} = (F_i [u_j]), \boldsymbol{\alpha}' = (\alpha_1, \alpha_2, \ldots, \alpha_m)', \mathbf{D} = \mathbf{diag}(d_1, \ldots, d_m),$$
$$\text{and } \mathbf{K} = (F_i [R_1 (\cdot, x_j)])$$

and its optimization is a matrix calculus problem. In the important special case where the F_i are function evaluation at x_i, above

$$F_i [u_j] = u_j (x_i) \text{ and } F_i [R_1 (\cdot, x_j)] = R_1 (x_i, x_j)$$