Green's function solution of splines

j.p.lewis

A general definition of a spline is *minimize the a weighted sum of squared derivatives, integrated over the (unknown) spline function, subject to passing through given points*, for example a natural cubic spline minimizes the integrated second derivative. This differential equation can be solved using the Green's function approach, which explains the form of RBF kernels. The cost is (cubically) proportional to the number of known/constrained points, rather than the number of unknown points (also cubically, if ignoring sparsity of the discretized derivative, and multigrid).

The general strategy is to use calculus of variations to setup a minimum for the energy, then use integration by parts to remove the derivatives from the variation test function $\epsilon h(x)$, express as $\int \epsilon h(x)(\cdots)dx$ and note that the $\cdots$ part must be zero. This converts the energy integral into a differential equation. Then we use a Green's function method: take the Fourier transform, solve for the function using simple algebra, take the inverse Fourier transform.

The following "energy" defines a thin plate with "tension":

$$F(f) = \sum (f(x_i) - y_i)^2 \quad + \quad \alpha \int \|\nabla f\|^2 dx \quad + \quad \gamma \int \|\nabla^2 f\|^2 dx$$

The summation covers the available data points, and the integration is over the domain of the spline.

An analogous one-dimensional problem is

$$F(f) = \sum (f(x_i) - y_i)^2 \quad + \quad \alpha \int \|\frac{df}{dx}\|^2 dx \quad + \quad \gamma \int \|\frac{d^2 f}{dx^2}\|^2 dx$$

Renotate this as

$$F(f) = \sum (f(x_i) - y_i)^2 \quad + \quad \alpha \int \|f_x\|^2 dx \quad + \quad \gamma \int \|f_{xx}\|^2 dx$$

The minimum of $F(f)$ can be found with variational calculus. In this approach, each occurrence of f_x in $F(f)$ is replaced with the variation $f_x + \epsilon h_x$. At a minimum, the variation

$$\frac{F(f + \epsilon h) - F(f)}{\epsilon}$$

will be zero. Put this in the form $\int hg$, then because h is arbitrary and the whole is zero, it must be that g is zero. To reduce notation we will temporarily call $f(x_i), h(x_i), y_i$ as f, h, y.

For the terms in the data-fit sum $\sum(f(x_i) - y_i)^2$,

$$
\begin{aligned}
&(f + \epsilon h - y)^2 \\
&= f^2 + e^2h^2 + y^2 \quad + \quad 2f\epsilon h - 2fy - 2\epsilon hy \\
&= f^2 - 2fy + y^2 \quad + \quad 2f\epsilon h - 2\epsilon hy + \epsilon^2 h^2 \\
&= (f - y)^2 + 2\epsilon h(f - y) + \epsilon^2 h^2
\end{aligned}
$$

The first derivative term $\alpha \int \|f_x\|^2$ expands to

$$
\begin{aligned}
&\alpha \int \|(f + \epsilon h)_x\|^2 dx \\
&= \alpha \int (f + \epsilon h)_x \quad (f + \epsilon h)_x dx \\
&= \alpha \int (f_x + \epsilon h_x)(f_x + \epsilon h_x) dx \\
&= \alpha \int \|f_x\|^2 + 2\epsilon f_x h_x + \epsilon^2 \|h_x\|^2 dx
\end{aligned}
$$

The second derivative term $\gamma \int \|f_{xx}\|^2 dx$ similarly expands to

$$
\begin{aligned}
&\gamma \int \|(f + \epsilon h)_{xx}\|^2 dx \\
&= \gamma \int (f + \epsilon h)_{xx} \quad (f + \epsilon h)_{xx} dx \\
&= \gamma \int (f_{xx} + \epsilon h_{xx})(f_{xx} + \epsilon h_{xx}) dx \\
&= \gamma \int \|f_{xx}\|^2 + 2\epsilon f_{xx} h_{xx} + \epsilon^2 \|h_{xx}\|^2 dx
\end{aligned}
$$

Now take the limit of

$$\frac{F(f + \epsilon h) - F(f)}{\epsilon}$$

as $\epsilon \to 0$. In dividing by ϵ, the single ϵ factors drop, whereas the ϵ^2 factors become ϵ. Then taking $\epsilon \to 0$ the ϵ becomes zero, effectively removing the ϵ^2 terms. In the following we will simply remove these terms as they appear.

$$
\begin{aligned}
&F(f+\epsilon h) - F(f) \setminus \{\epsilon^2 \text{ terms}\} = \\
&\sum (f(x_i)-y_i)^2 + 2\epsilon h(x_i)(f(x_i)-y_i) \\
&\quad + \quad \alpha \int \|f_x\|^2 + 2\epsilon f_x h_x \quad dx \\
&\quad + \quad \gamma \int \|f_{xx}\|^2 + 2\epsilon f_{xx} h_{xx} \quad dx \\
&- \quad \sum (f(x_i)-y_i)^2 \quad - \quad \alpha \int \|f_x\|^2 dx \quad - \quad \gamma \int \|f_{xx}\|^2 dx \\
&= 2\epsilon \sum h(x_i)(f(x_i)-y_i) \quad + \quad 2\epsilon\alpha \int f_x h_x dx \quad + \quad 2\epsilon\gamma \int f_{xx} h_{xx} dx
\end{aligned}
$$

Pull the sum under an integral by expressing it with a Dirac:

$$
= 2 \int \epsilon \sum h(x)(f(x_i)-y_i)\delta(x-x_i) dx \quad + \quad 2\epsilon\alpha \int f_x h_x dx \quad + \quad 2\epsilon\gamma \int f_{xx} h_{xx} dx
$$

Now do two integration by parts (IBP) on the second derivative term, thereby removing the derivative from h and increasing it on f. Likewise do one IBP on the first derivative term. Recall the integration by parts pattern:

$$
\int AB_x dx = AB|_b^e - \int BA_x dx
$$

where A_x is the derivative of some function $A(x)$, and b, e denote the beginning and end of the domain of integration.

The result of applying IBP to the first derivative term is:

$$
2\epsilon\alpha \int f_x h_x dx = -2\epsilon\alpha \int h(x) f_{xx}(x) dx
$$

The AB term in the IBP is zero because the test function $h(x)$ is zero at the boundaries.

Second derivative IBP

$$
\begin{aligned}
& 2 \int \epsilon f_{xx} h_{xx} dx \\
\text{first application of IBP: } & = -2\epsilon\gamma \int f_{xxx} h_x dx \\
\text{second application of IBP: } & = 2\epsilon\gamma \int f_{xxxx} h dx
\end{aligned}
$$

Substituting these IBP expressions, the new $F(f+\epsilon h) - F(f)$ is

$$
2\epsilon \int \sum h(x)(f(x_i)-y_i)\delta(x-x_i) dx \quad - \quad 2\epsilon\alpha \int h(x) f_{xx}(x) dx \quad + \quad 2\epsilon\gamma \int f_{xxxx} h dx
$$

$$= 2 \int \epsilon h(x) \left(\sum (f(x_i) - y_i) \delta(x - x_i) dx \quad - \quad \alpha f_{xx}(x) \quad + \quad \gamma f_{xxxx} \right) dx$$

and finally taking $\frac{F(f+\epsilon h)-F(f)}{\epsilon}$ eliminates the ϵ. This expression is zero at the minimum because it is the variation. Then, because $h(x)$ can be arbitrary, we know that the quantity in parenthesis is zero everywhere:

$$\gamma f_{xxxx}(x) \quad - \quad \alpha f_{xx}(x) \quad + \quad \sum (f(x_i) - y_i) \delta(x - x_i) = 0 \tag{1}$$

(Note that the minus sign came from the IBP).

This differential equation can be solved with the Green's function approach. For a Green's function setup,

$$\gamma g_{xxxx}(x) - \alpha g_{xx}(x) = \delta(x) \tag{2}$$

and the function is expressed as a weighted convolution of the Green's function with the data:

$$f(x) = \sum_k w_k g(x - x_k) \tag{3}$$

Substitute this definition into Eq. (1) (repeated here):

$$\begin{aligned}
&\gamma f_{xxxx}(x) - \alpha f_{xx}(x) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0 \\
=&\gamma \frac{d^4}{dx^4} (\sum_k w_k g(x - x_k)) - \alpha \frac{d^2}{dx^2} (\sum_k w_k g(x - x_k)) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0 \\
=&\gamma \sum w_k \frac{d^4}{dx^4} g(x - x_k) - \alpha \sum w_k \frac{d^2}{dx^2} g(x - x_k) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0
\end{aligned}$$

(Changing notation)

$$\begin{aligned}
=&\gamma \sum w_k g_{xxxx}(x - x_k) - \alpha \sum w_k g_{xx}(x - x_k) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0 \\
=&\sum w_k \gamma g_{xxxx}(x - x_k) - \sum w_k \alpha g_{xx}(x - x_k) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0 \\
=&\sum w_k \left(\gamma g_{xxxx}(x - x_k) - \alpha g_{xx}(x - x_k) \right) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0 \\
=&\sum w_k \delta(x - x_k) + \sum_i (f(x_i) - y_i) \delta(x - x_i) = 0
\end{aligned}$$

Where the last step comes from applying Eq. (2).

Now substitute the expression Eq. (3) for $f(x)$ in terms of the Green's function on the right hand side, giving

$$\sum w_i \delta(x - x_i) + \sum_i ((\sum_k w_k g(x_i - x_k)) - y_i) \delta(x - x_i) = 0$$

Note the switch from k to i in the leftmost term. This is ok – the x_k on the left indexes the same data points as x_i on the right.

The $\delta(x - x_i)$ are independent, that is,

$$\sum a_k \delta(x - x_k) = 0 \quad \Rightarrow \quad a_k = 0$$

so this means

$$w_i + ((\sum_k w_k g(x_i - x_k)) - y_i) = 0 \qquad \forall i$$

or

$$\sum_k w_k g(x_i - x_k) + w_i - y_i = 0 \qquad \forall i$$

which resembles is a linear system for the approximation problem... but the function g is still unknown! What happened to α, γ? They are "baked" into the Greens function.

To solve for the Greens function g: Go back to the Greens function relation (2) and take the Fourier transform,

$$\gamma j^4 \omega^4 G(\omega) - \alpha j^2 \omega^2 G(\omega) = 1$$

$$G(\omega) = \frac{1}{\gamma j^4 \omega^4 - \alpha j^2 \omega^2}$$

Since $j^2 = -1$ and $j^4 = 1$,

$$G(\omega) = \frac{1}{\gamma \omega^4 + \alpha \omega^2}$$

To simplify, we will take $\gamma = 1$. Then this is equal to

$$\frac{1}{\omega^2} \frac{1}{\alpha + \omega^2}$$

The $\frac{1}{\alpha+\omega^2}$ is the transform of the Laplace kernel, so the desired kernel is (any) second antiderivative of the Laplace kernel.

Note that since the original roughness penalty does not "see" constants, a polynomial correction of order 0 (i.e. a constant) is necessary.

Acknowledgements: I learned this from Jean-Marie Aubrey and Jan Kybic's thesis.