Reproducing Kernel Hilbert Space

November 11, 2010

Say we want a nice space of functions. First choose a symmetric, positive
definite kernel \( k(x, y) \). Each point defines a basis function

\[
k(x, \cdot).
\]

(1)

Now make a vector space of functions (RKHS)

\[
f(x) = \sum_{i=1}^{n} \alpha_i k(x_i, x).
\]

(2)

Define the inner product as

\[
\langle f, g \rangle = \sum_{i=1}^{n} \sum_{m} \sum_{j=1}^{m} \alpha_i \beta_j k(x_i, y_j).
\]

(3)

This inner product is valid because \( k \) is positive definite.

The kernel is like the Dirac delta function (representer of evaluation):

\[
\langle f, k(x, \cdot) \rangle = \sum_{i=1}^{n} \alpha_i k(x_i, x) = f(x).
\]

(4)

The kernel is the inner product for basis functions (reproducing property, kernel
trick):

\[
\langle k(x, \cdot), k(y, \cdot) \rangle = k(x, y).
\]

(5)

The induced norm is

\[
\|f\| = \langle f, f \rangle = \sum_{i, j=1}^{n} \alpha_i^2 k(x_i, x_j).
\]

(6)

Say we want to minimize

\[
J = \sum_{i} (f(x_i) - y_i)^2 + \lambda \|f\|^2.
\]

(7)

The representer theorem says that the optimal \( f \) is

\[
f(x) = \sum_{i} \alpha_i k(x_i, x).
\]

(8)
Proof.

\[ \hat{f}(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + f_\perp(x), \]  

(9)

\[ \hat{f}(x_j) = \sum_{i=1}^{n} \alpha_i K(x_i, x_j), \]  

(10)

\[ \| \hat{f} \|^2 = \| \sum_{i=1}^{n} \alpha_i K(x_i, \cdot) \|^2 + \| f_\perp \|^2, \]  

(11)

\[ \hat{f} = f. \]  

(12)

The Gaussian kernel is

\[ k(x, y) = e^{-\| x - y \|^2}. \]  

(13)

It maps points to the first quadrant of the sphere because

\[ \| k(x, \cdot) \|^2 = k(x, x) = 1 \]  

(14)

\[ \cos(\alpha_{x,y}) = k(x, y) \geq 0. \]  

(15)