

9.520 Problem set 2

Pr. 2.1 Consider the following one dimensional RBF *interpolation* (ie $\lambda = 0$ in the regularization network formulation) scheme:

$$f(x) = \sum_{i=1}^{\ell} c_i K(x - x_i), \quad (1)$$

with $K(x) = |x|$ and $\ell = 3$.

1. Compute the coefficients c_i , $i = 1, 2, 3$, in such a way that the following interpolation conditions are satisfied:

$$f(0) = 1, \quad f(3) = 4, \quad f(4) = 3$$

and draw the corresponding curve.

2. Show that, on the interval $[0, 4]$, the Radial Basis Functions expansion (1) can be also written in the form

$$f(x) = \sum_{i=1}^{\ell} y_i b_i(x)$$

where the y_i are the values to be interpolated ($y_1 = 1$, $y_2 = 4$, and $y_3 = 3$). Derive and draw the explicit form for the dual kernels $b_i(x)$.

Pr. 2.2 Consider the following variational problem:

$$\min_{f \in \mathcal{H}} \frac{1}{\ell} \sum_{i=1}^{\ell} \left(\int_{-\infty}^{x_i} f(t) dt - F(x_i) \right)^2 + \lambda \int_{-\infty}^{\infty} \frac{\hat{f}(\omega) \hat{f}(\omega)}{\hat{K}(\omega)}, \quad (2)$$

where $\hat{f}(\omega)$ is the Fourier transform of $f(x)$, $\hat{K}(\omega)$ is the Fourier transform of the kernel, and $F(x_i)$ is the empirical cumulative distribution function (cdf). Assume the kernel to be a Gaussian, $K(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-y)^2/2\sigma^2}$.

Show that the solution of the above problem has

$$f(x) = \frac{1}{\ell} \sum_{i=1}^{\ell} K(x, x_i).$$

Compare this solution to the density estimator used in the Parzen's windows algorithm.

Pr. 2.3 Prove that the solution to equation (2) has stability of order $\frac{1}{\ell}$. Is this always the case for the above variational problem or are further restrictions required to get stability of order $\frac{1}{\ell}$?

Pr 2.4 Let S_n be the set of all the hyperplanes in \mathbb{R}^n .

1. Show that S_2 separates any three points not lying on the same line in \mathbb{R}^2 in all possible ways;
2. Show that S_2 cannot separate any four points of \mathbb{R}^2 in all possible ways.
3. What is the minimal number of points in \mathbb{R}^{100000} that cannot be separated in all possible ways by S_{100000} ?

Pr 2.5 Find the VC dimension of conics for points in the plane. Is there any difference if you restrict the set to one type of conics only (ellipses, parabolae, and hyperbolae)? What happens in the case of degenerate conics?

Pr 2.6 Assuming that the generalization error has the form

$$E(n, \ell) \leq \frac{1}{n^\alpha} + \sqrt{\frac{n^\beta}{\ell}}$$

determine the optimal number of parameters, n , as a function of the number of examples, ℓ , and estimate the best possible rate of convergence.