Abstract

Motivated by the classic works of Charles M. Stein, we focus on developing risk-estimation frameworks for denoising problems in both one- and two-dimensions. We assume a standard additive noise model, and formulate the denoising problem as one of estimating the underlying clean signal from noisy measurements by minimizing a risk corresponding to a chosen loss function. Our goal is to incorporate perceptually-motivated loss functions wherever applicable, as in the case of speech enhancement, with the squared error loss being considered for the other scenarios. Since the true risks are observed to depend on the unknown parameter of interest, we circumvent the roadblock by deriving finite-sample unbiased estimators of the corresponding risks based on Stein's lemma. We establish the link with the multivariate parameter estimation problem addressed by Stein and our denoising problem, and derive estimators of the oracle risks. In all cases, optimum values of the parameters characterizing the denoising algorithm are determined by minimizing the Stein's unbiased risk estimator (SURE).

The key contribution of this thesis is the development of a risk-estimation approach for choosing the two critical parameters affecting the quality of nonparametric regression, namely, the order and bandwidth/smoothing parameters. This is a classic problem in statistics, and certain algorithms relying on derivation of suitable finite-sample risk estimators for minimization have been reported in the literature (note that all these works consider the mean squared error (MSE) objective). We show that a SURE-based formalism is well-suited to the regression parameter selection problem, and that the optimum solution guarantees near-minimum MSE (MMSE) performance. We develop algorithms for both globally and locally choosing the two parameters, the latter referred to as spatially-adaptive regression. We observe that the parameters are so chosen as to tradeoff the squared bias and variance quantities that constitute the MSE. We also indicate the advantages accruing out of incorporating a regularization term in the cost function in addition to the data error term. In the more general case of kernel regression, which uses a weighted least-squares (LS) optimization, we consider the applications of image restoration from very few random measurements, in addition to denoising of uniformly sampled data. We show that local polynomial regression (LPR) becomes a special case of kernel regression, and extend our results for LPR on uniform data to non-uniformly sampled data also. The denoising algorithms are compared with other standard, performant methods available in the literature

both in terms of estimation error and computational complexity.

A major perspective provided in this thesis is that the problem of optimum parameter choice in nonparametric regression can be viewed as the selection of optimum parameters of a linear, shift-invariant filter. This interpretation is provided by deriving motivation out of the hallmark paper of Savitzky and Golay and Schafer's recent article in IEEE Signal Processing Magazine. It is worth noting that Savitzky and Golay had shown in their original Analytical Chemistry journal article, that LS fitting of a fixed-order polynomial over a neighborhood of fixed size is equivalent to convolution with an impulse response that is *fixed* and can be *pre-computed*. They had provided tables of impulse response coefficients for computing the smoothed function and smoothed derivatives for different orders and neighborhood sizes, the resulting filters being referred to as Savitzky-Golay (S-G) filters. Thus, we provide the new perspective that the regression parameter choice is equivalent to optimizing for the filter impulse response length/3 dB bandwidth, which are inversely related. We observe that the MMSE solution is such that the S-G filter chosen is of longer impulse response length (equivalently smaller cutoff frequency) at relatively flat portions of the noisy signal so as to smooth noise, and vice versa at locally fast-varying portions of the signal so as to capture the signal patterns. Also, we provide a generalized S-G filtering viewpoint in the case of kernel regression.

Building on the S-G filtering perspective, we turn to the problem of dynamic feature computation in speech recognition. We observe that the methodology employed for computing dynamic features from the trajectories of static features is in fact derivative S-G filtering. With this perspective, we note that the filter coefficients can be pre-computed, and that the whole problem of delta feature computation becomes efficient. Indeed, we observe an advantage by a factor of 10^4 on making use of S-G filtering over actual LS polynomial fitting and evaluation. Thereafter, we study the properties of first- and second-order derivative S-G filters of certain orders and lengths experimentally. The derivative filters are bandpass due to the combined effects of LPR and derivative computation, which are lowpass and highpass operations, respectively. The first- and second-order S-G derivative filters are also observed to exhibit an approximately constant-Q property. We perform a TIMIT phoneme recognition experiment comparing the recognition accuracies obtained using S-G filters and the conventional approach followed in HTK, where Furui's regression formula is made use of. The recognition accuracies for both cases are almost identical, with S-G filters of certain bandwidths and orders registering a marginal improvement. The accuracies are also observed to improve with longer filter lengths, for a particular order. In terms of computation latency, we note that S-G filtering achieves delta and delta-delta feature computation in parallel by linear filtering, whereas they need to be obtained sequentially in case of the standard regression formulas used in the literature.

Finally, we turn to the problem of speech enhancement where we are interested in denoising using perceptually-motivated loss functions such as Itakura-Saito (IS). We propose to perform enhancement in the discrete cosine transform domain using risk-minimization. The cost functions considered are non-quadratic, and derivation of the unbiased estimator of the risk corresponding to the IS distortion is achieved using an approximate Taylor-series

analysis under high signal-to-noise ratio assumption. The exposition is general since we focus on an additive noise model with the noise density assumed to fall within the exponential class of density functions, which comprises most of the common densities. The denoising function is assumed to be pointwise linear (modified James-Stein (MJS) estimator), and parallels between Wiener filtering and the optimum MJS estimator are discussed.