Eigenfunction methods for estimation with random fields

Christopher G. Small

Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
[cgsmall@uwaterloo.ca]

It is a well accepted that optimal inference about the mean of a distribution requires some knowledge of its dispersion and shape. For a continuous-time stochastic process or for a random field, fully optimal inference about the mean function requires knowledge about the **covariance kernel**, which is the limiting form of the covariance matrix for data in finite dimensions. Suppose that $X(\mathbf{t})$ is a real-valued stochastic process or field, where $\mathbf{t} \in I\!R^q$. In practice $X(\mathbf{t})$ will only be observed for $\mathbf{t}$ within some bounded "window." For the purposes of this paper, we shall assume that $\mathbf{t}$ lies in some bounded closed subset $\mathcal{R}$ of $I\!R^q$ with nonempty interior.

$$E_\theta[X(\mathbf{t})] = \mu_\theta(\mathbf{t}) \text{ and } \quad \mathrm{Cov}_\theta[X(\mathbf{s}), X(\mathbf{t})] = \delta\, \Gamma_\theta(\mathbf{s}, \mathbf{t})$$

be the mean function and covariance kernel respectively, where $\mathbf{s}$, $\mathbf{t} \in \mathcal{R}$. We let θ denote a k-dimensional column vector of parameters. Both μ_θ and Γ_θ are assumed to be known real-valued functions of the unknown parameter θ. The quantity $\delta > 0$ is assumed to be unknown, or known and equal to to one for some models, but functionally free of θ.

To estimate θ or $\mu(\theta)$, the Karhunen-Loève expansion provides a useful decomposition of the process into eigenfunction for the covariance kernel. This leads to an estimating equation for θ of the form

$$\sum_{j=1}^{\infty} \rho_j^{-1}(\theta)\tau_j(\theta)[Y_j(\theta) - \nu_j(\theta)] = 0,$$

where ρ_j, τ_j and ν_j arise from the K.-L. expansion, and play the role of the mean gradient, the variance, and the mean of $Y - j$ respective. The formulation reduces to the classic quasi-likelihood (generalised estimating equation) setting.

This elegant solution to the estimation problem masks a number of practical and theoretic difficulties. For many data sets the covariance kernel must be estimated. But this is problematic if stationarity cannot be assumed. Secondly, the actual eigenfunction decomposition is nontrivial to compute for many kernels. In this talk, I shall consider a family of "working kernels" analogous to the working covariance matrices of quasi-likelihood or generalised estimating equations.