

Date of Submission (month day, year) : July 6, 2026

Department of Computer Science and Engineering	Student ID Number D249302	Supervisors Kazuho Watanabe Norihide Kitaoka
Applicant's name Tsukasa Yoshida		

Abstract (Doctor)

Title of Thesis	Empirical Bayes Estimation for Regularization Parameters in Sparse Modeling
-----------------	---

Approx. 800 words

In recent years, the collection of large-scale datasets has become increasingly feasible. However, truly informative components are often confined to a small subset, making the extraction of essential information from massive data a fundamental challenge. Methods that assume sparsity in the underlying system are collectively known as sparse modeling. In regression problems, Lasso is a representative approach for feature selection under the sparsity assumption, which is achieved by adding an l_1 penalty to the loss function of a linear regression model. Several extensions, including Elastic Net, Group Lasso, Fused Lasso, and Total Variation regularization, have also been extensively studied. The performance of these methods depends critically on the regularization parameter, which controls the degree of sparsity imposed on the estimated coefficients. A parameter value that is too small fails to achieve sufficient sparsity and leads to overfitting, while an excessively large value suppresses too many coefficients and causes loss of relevant information. Although effective, these methods are highly sensitive to the choice of the regularization parameter, which is typically selected via cross-validation at considerable computational cost, especially in high-dimensional settings. To address this issue, we focus on the empirical Bayes framework, in which the regularization parameter is estimated through maximum marginal likelihood. This approach treats the regularization parameter as a hyperparameter and determines it directly from the data, providing a statistically principled alternative to cross-validation. Although empirical Bayes is supported by a well-established mathematical foundation, the likelihood and prior distribution are not conjugate when regularization methods such as Lasso are employed. For Ridge regression, the Gaussian prior on the regression coefficients is conjugate to the Gaussian likelihood, which allows the marginal likelihood to be computed in closed form; the resulting empirical Bayes estimator has been characterized in prior work. In contrast, Lasso and related methods correspond to a Laplace prior, for which no such conjugacy holds. Consequently, the marginal likelihood involves an analytically intractable integral. As a result, general closed-form solutions of the empirical Bayes estimator and their theoretical properties remain largely unexplored. To circumvent this difficulty, approximate inference methods such as variational Bayes have been proposed. However, since these approaches rely on iterative
--

optimization procedures and yield only approximate solutions, the structural properties and precise behavior of the resulting empirical Bayes estimators have not yet been fully clarified.

The objective of this study is to derive analytical solutions for the empirical Bayes estimation of regularization parameters in sparse modeling and to rigorously characterize their theoretical properties. In particular, we focus on regularization methods possessing homogeneity properties, such as Ridge, Lasso, and Group Lasso. We conduct a theoretical analysis of regularized linear regression models under this framework. First, we prove that the marginal likelihood of these models with a certain dimension of the regression parameter is quasi-concave. This quasi-concavity is the key to our analysis: it guarantees that the empirical Bayes estimator is uniquely determined and enables its exact closed-form characterization. Based on this property, we derive the exact empirical Bayes estimator and establish an explicit upper bound on its value. We further show that the estimator exhibits a phase transition, with its qualitative behavior changing beyond a threshold. Notably, this threshold is found to be identical across Ridge, Lasso, and Group Lasso, despite their different regularization structures, revealing a universal phase transition structure shared by these methods. We then analyze the maximum a posteriori (MAP) estimator with the empirical Bayes estimate substituted for the regularization parameter, and characterize the resulting thresholding function in detail. Moreover, we extend the analysis to a general class of homogeneous regularized linear regression models and derive the empirical Bayes estimator under the quasi-concavity condition of the marginal likelihood. In this generalized setting, we again demonstrate the occurrence of a phase transition and explicitly characterize its condition. Specifically, when the observed signal is weak relative to the noise level, the estimator diverges to infinity and forces the corresponding MAP coefficient to exactly zero; when the signal is sufficiently strong, a finite estimate is obtained and the coefficient is estimated as nonzero. The phase transition condition can thus be interpreted as a criterion for signal detectability: it specifies whether a signal embedded in observation noise is strong enough to be identified from the data. Finally, we study the thresholding function of the MAP estimator based on the empirical Bayes estimate and show its similarity to Adaptive Lasso. We prove that MAP estimation under the empirical Bayes estimate of the regularization parameter does not satisfy variable selection consistency in general; nonetheless, all variables with nonzero true coefficients are always selected asymptotically. In other words, when the sample size is sufficiently large, no variable in the true support (i.e., the set of indices with nonzero true coefficients) is ever estimated as zero.