Student Research Symposium 2004

Guy Lebanon

Hyperplane Margin Classifiers on the Multinomial Manifold

Linear classifiers are a mainstay of machine learning algorithms, forming the basis for techniques such as the perceptron, logistic regression, boosting, and support vector machines. A linear classifier, parameterized by a vector $w \in \mathbb{R}^n$, classifies examples according to the decision rule

$$\hat{y}(x) = \text{sign}\left(\sum_i w_i \phi_i(x)\right) = \text{sign}(\langle w, x \rangle) \in \{-1, +1\},$$

following the common practice of identifying x with the feature vector $\phi(x)$. The differences between different linear classifiers lie in the criteria and algorithms used for selecting the parameter vector w based on a training set.

Geometrically, the decision surface of a linear classifier is formed by a hyperplane or linear subspace in n-dimensional Euclidean space, $\{x \in \mathbb{R}^n : \langle x, w \rangle = 0\}$ where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. (In both the algebraic and geometric formulations, a bias term is sometimes added; we prefer to absorb the bias into the notation given by the inner product, by setting $x_n = 1$ for all x.) The linearity assumption made by such classifiers may be justified by its solution to the fundamental learning tradeoff between complexity of models and restricted expressiveness.

However, we show that implicit in this argument is the presence of Euclidean geometry. If the data is not well described by Euclidean geometry, the main motivation for linear classifiers fails and a generalization of linear classifiers, adapted to the present geometry is expected to perform better. In this work, we generalize the notion of linear hyperplane and margin to arbitrary Riemannian geometries. The natural generalization of logistic regression is then defined and its properties are examined. We focus our attention on the Fisher geometry of the multinomial manifold that forms a natural geometric space for text documents. The resulting generalization of logistic regression is shown to outperform its Euclidean counterpart on several standard text classification tasks.

Relevant Publication:

G. Lebanon and J. Lafferty. Hyperplane Margin Classifiers on the Multinomial Manifold. *International Conference on Machine Learning* (ICML) 2004.