COMP4670/6467
Introduction to Statistical Machine Learning
Tutorial 3

Christfried Webers

19/21 March 2013

1 Classification via Logistic Regression

Implement classification via logistic regression for a data set with two classes (as discussed in the lecture). Use the Fisher Iris data (available from the course web site) and assume for this exercise that Iris-setosa is class $\mathcal{C}_1$ and both Iris-versicolor and Iris-verginica belong to class $\mathcal{C}_2$.

More information about the Fisher Iris data set can be found under `http://en.wikipedia.org/wiki/Iris_flower_data_set`.

2 Marginal and Conditional Gaussians

In the lecture we stated the following result.

- The Gaussian family is conjugate to itself with respect to a Gaussian likelihood function: if the likelihood function is Gaussian, choosing a Gaussian prior will ensure that the posterior distribution is also Gaussian.
- Given a marginal distribution for $\mathbf{x}$ and a conditional Gaussian distribution for $\mathbf{y}$ given $\mathbf{x}$ in the form

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x} \,|\, \boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1})$$
$$p(\mathbf{y} \,|\, \mathbf{x}) = \mathcal{N}(\mathbf{y} \,|\, \boldsymbol{A}x + \mathbf{b}, \boldsymbol{L}^{-1})$$

- we get

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{y} \,|\, \boldsymbol{A}\boldsymbol{\mu} + \mathbf{b}, \boldsymbol{L}^{-1} + \boldsymbol{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{A}^T)$$
$$p(\mathbf{x} \,|\, \mathbf{y}) = \mathcal{N}(\mathbf{x} \,|\, \boldsymbol{\Sigma}\{\boldsymbol{A}^T\boldsymbol{L}(\mathbf{y} - \mathbf{b}) + \boldsymbol{\Lambda}\boldsymbol{\mu}\}, \boldsymbol{\Sigma})$$

where $\boldsymbol{\Sigma} = (\boldsymbol{\Lambda} + \boldsymbol{A}^T\boldsymbol{L}\boldsymbol{A})^{-1}$.

Prove the result for $p(\mathbf{y})$ and $p(\mathbf{x} \mid \mathbf{y})$ given $p(\mathbf{x})$ and $p(\mathbf{y} \mid \mathbf{x})$. (As a first step, it may help to assume that $\mathbf{x}$ and $\mathbf{y}$ are scalar. Then extend the result to vectors.)