Advanced Statistical Inference
Bayesian Logistic Regression - Metropolis-Hastings

1 Aims

- To implement the MH algorithm.
- To use it to compute classification probabilities.

2 Metropolis-Hastings

In this lab, you're going to implement the Metropolis-Hasting algorithm described in the lecture. Use the binary classification data `binaryclass2.mat` and the function `laplacecomp.m`. If you pass this function a 2-dimensional $\mathbf{w}$ vector, it will return $g(\mathbf{w}; \mathbf{X}, \mathbf{t}, \sigma^2)$ and $\log g(\mathbf{w}; \mathbf{X}, \mathbf{t}, \sigma^2)$. (Remember that $g(\mathbf{w}; \mathbf{X}, \mathbf{t}, \sigma^2) \propto p(\mathbf{w}|\mathbf{X}, \mathbf{t}, \sigma^2)$, the posterior density of interest.)

1. Make sure `laplacecomp.m` works by passing it some different $\mathbf{w}$ vectors ($\mathbf{w}$ needs to have dimension 2×1).

2. Implement the MH algorithm (see the flowchart in the slides). You might find it safest to work with $\log g$ – the ratio then becomes a subtraction that must then be exponentiated. This has more numerical stability, especially when $\mathbf{w}$ isn't very good. You can start with any value of $\mathbf{w}$. e.g. `w = randn(2,1)`. Also, use a Gaussian proposal: For example: `wp=randn(2,1)*0.5+w` where `w` is the current sample and `wp` is the proposal and the Gaussian has standard deviation 0.5 (set this as you please).

3. Compute the probability that $P(t_{\mathsf{new}} = 1|\mathbf{x}_{\mathsf{new}}, \mathbf{X}, \mathbf{t})$ when $\mathbf{x}_{\mathsf{new}} = [2, -4]^{\mathsf{T}}$. Hint – compute the probability for each value of $\mathbf{w}$ using:

$$\frac{1}{1 + \exp(-\mathbf{w}^{\mathsf{T}}\mathbf{x}_{\mathsf{new}})}$$

   and average!