Computing logistic regression predictions

In the previous note we approximated the logistic regression posterior with a Gaussian distribution. By comparing to the joint probability, we immediately obtained an approximation for the marginal likelihood $P(D)$ or $P(D|M)$, which can be used to choose between alternative model settings $M$.

Now we return to the question of how to make Bayesian predictions (all implicitly conditioned on a set of model choices $M$):

\[
P(y|x,D) = \int p(y,w|x,D) \, dw = \int P(y|x,w) \, p(w|D) \, dw.
\]

If we approximate the posterior with a Gaussian, $p(w|D) \approx \mathcal{N}(w;m,V)$, we still have an integral with no closed form solution:

\[
P(y=1|x,D) \approx \int \sigma(w^\top x) \, \mathcal{N}(w;m,V) \, dw = E_{\mathcal{N}(w;m,V)}[\sigma(w^\top x)].
\]

However, this expectation can be simplified. Only the inner product $a = w^\top x$ matters, so we can take the average over this scalar quantity instead. The linear combination $a$ is a linear combination of Gaussian beliefs, so our beliefs about it are also Gaussian. By now you should be able to show that

\[
p(a) = \mathcal{N}(a; m^\top x, x^\top Vx).
\]

Therefore, the predictions given the approximate posterior, are given by a one-dimensional integral:

\[
P(y=1|x,D) \approx E_{\mathcal{N}(a;m^\top x, x^\top Vx)}[\sigma(a)] = \int \sigma(a) \, \mathcal{N}(a; m^\top x, x^\top Vx) \, da.
\]

One-dimensional integrals can be computed numerically to high precision.

Murphy Section 8.4.4.2 reviews a further approximation (derivation non-examinable), which results in a closed form expression:

\[
P(y=1|x,D) \approx \sigma(\kappa m^\top x), \quad \kappa = \frac{1}{\sqrt{1 + \frac{a}{x^\top Vx}}}.
\]

These predictions use the mean weights $m$ under the Gaussian approximation. If we used the Laplace approximation, we’re using the most probable or MAP weights. However, the activation is scaled down (with $\kappa$) when the activation is uncertain, so that predictions will be less confident far from the data (as they should be).