Bayesian Learning/Inference

- Recall simple univariate Gaussian example (Ch 4)

MLPs – Bayesian Learning

- Section 11.10.
- Bayesian learning can be applied to the task of training an MLP (finding weight values).
- Slide in notes is assuming independent weights and a prior (p(w) Gaussian with mean 0 and variance ½*lambda.
- Requires approximation or MCMC simulation to implement.
  - But apparently performs really well!
- Again see book by Bishop.

Bayesian Learning

- Consider weights w_i as random vars, prior $p(w_i)$

$$p(\mathbf{w} \mid \mathcal{X}) = \frac{p(\mathcal{X} \mid \mathbf{w})p(\mathbf{w})}{p(\mathcal{X})} \quad \hat{\mathbf{w}}_{MAP} = \arg\max_{\mathbf{w}} \log p(\mathbf{w} \mid \mathcal{X})$$

$$\log p(\mathbf{w} \mid \mathcal{X}) = \log p(\mathcal{X} \mid \mathbf{w}) + \log p(\mathbf{w}) + C$$

$$p(\mathbf{w}) = \prod_i p(w_i) \text{ where } p(w_i) = c \cdot \exp\left[-\frac{w_i^2}{2(1/2\lambda)}\right]$$

$$E' = E + \lambda \|\mathbf{w}\|^2$$

- Weight decay, ridge regression, regularization

  cost=data-misfit + λ complexity

Combining learners (Chap 15)

- Bayesian perspective (in 15.2)
- Bayesian methodology suggest predict using the posterior over models, not just a single model.
  - In practice this can be done with a sample of models from the posterior (1st eqn on slide 5, or eqn 15.4 in text).
  - Becomes an ensemble/voting scheme. Simple voting = uniform prior.
  - Shown to decrease variance of model error, without increasing bias.

- Bayesian perspective:

$$P(C_i \mid x) = \sum_{\text{all models } \mathcal{M}_j} P(C_i \mid x, \mathcal{M}_j) P(\mathcal{M}_j)$$

- If d_j are iid

$$E[y] = E\left[\sum_j \frac{1}{L} d_j\right] = \frac{1}{L} L \cdot E[d_j] = E[d_j]$$

$$\text{Var}(y) = \text{Var}\left(\sum_j \frac{1}{L} d_j\right) = \frac{1}{L^2} \text{Var}\left(\sum_j d_j\right) = \frac{1}{L^2} L \cdot \text{Var}(d_j) = \frac{1}{L} \text{Var}(d_j)$$

  Bias does not change, variance decreases by L

- Average over randomness

Where to from here?

- Thesis projects in AI - more specifically
  - Machine Learning
  - Nature-inspired and metaheuristic optimization
  - Cool applications of these.
- Check the projectdb for some projects before start of next semester, or come and chat to me sometime ☺

Where to from there?

- Postgrad research (PhD, MPhil) in same areas:
  - Opens up interesting career paths (really!)
  - Some tax-free scholarships available (with good GPA).
  - Happy to talk to you anytime ☺