Lecture 15: Advanced topics in Bayesian linear regression

Professor Ilias Bilionis

Automatic relevance determination

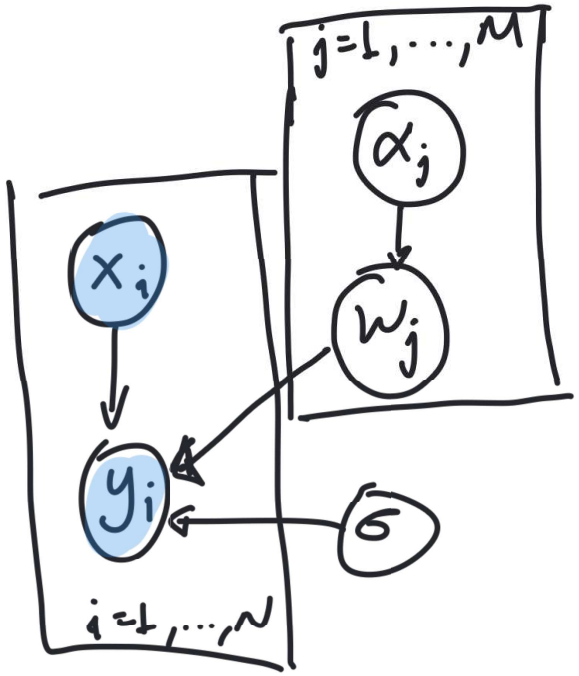
Open questions

- How do I quantify the measurement noise?
- How do we avoid overfitting?
- How do I quantify epistemic uncertainty induced by limited data?

Lecture 14

- How do I choose any remaining parameters? ✓
- How do I choose which basis functions to keep?
  ↳ automatic relevance detection

Idea: Different hyper-prior per weight



Prior:

$$\alpha_j \sim p(\alpha_j)$$

$$w_j|\alpha_j \sim p(w_j|\alpha_j) = N(w_j|0,\alpha_j^{-1})$$

zero-mean with precision/variance determined by α

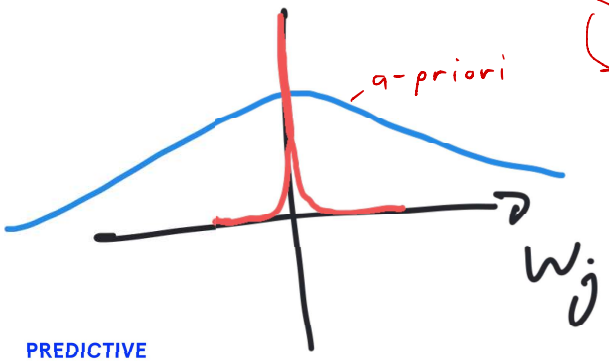
$$\sigma \sim p(\sigma)$$

Likelihood:

evidence approximation: $\underline{\alpha}^*, \sigma^* = \arg\max_{\underline{\alpha},\sigma} p(\underline{\alpha}, \sigma | x_{1:N}, y_{1:N})$ ← marginalized posterior

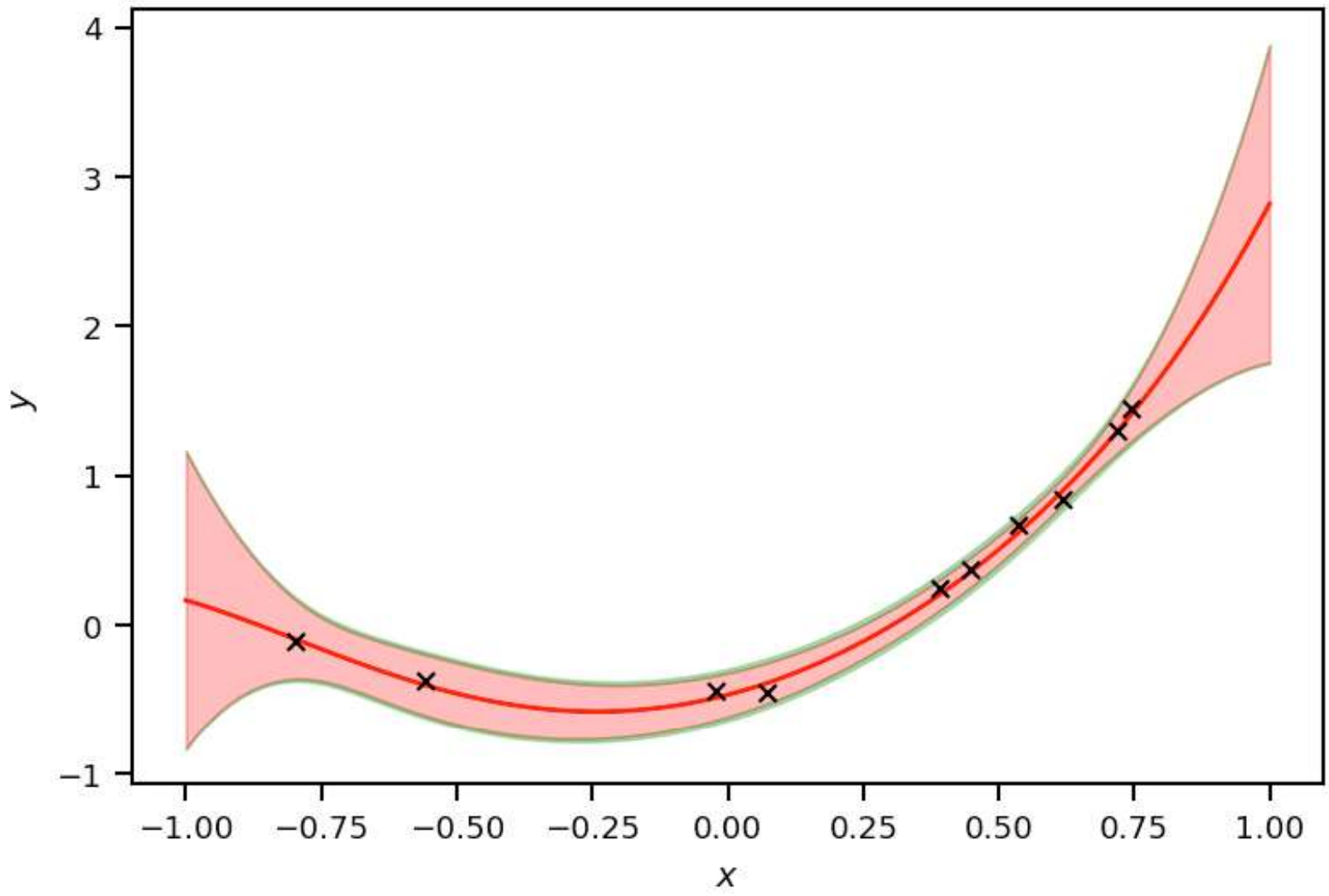
finding max α's

$\alpha_j >> \infty \rightarrow w_j \rightarrow 0$

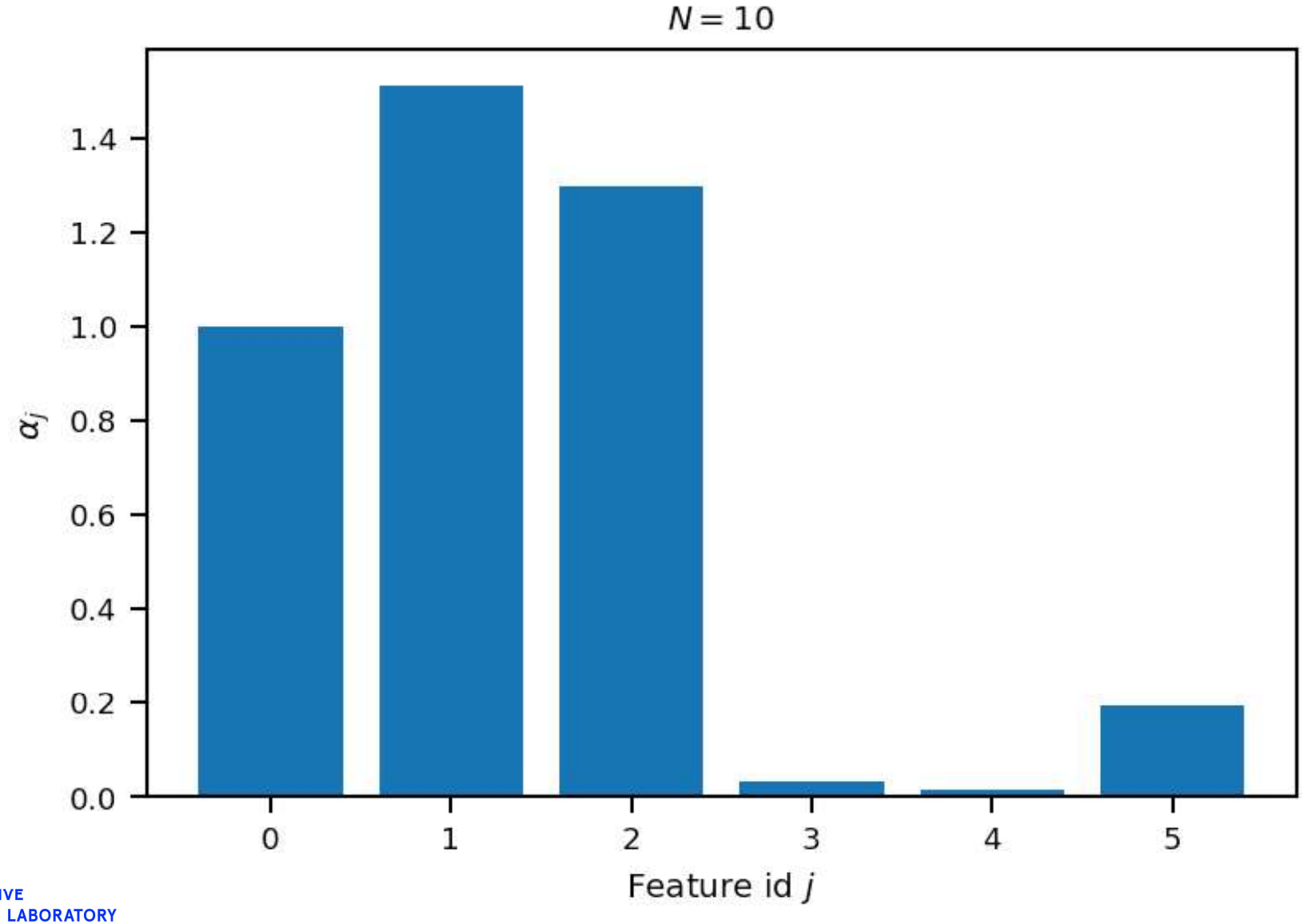


PREDICTIVE SCIENCE LABORATORY

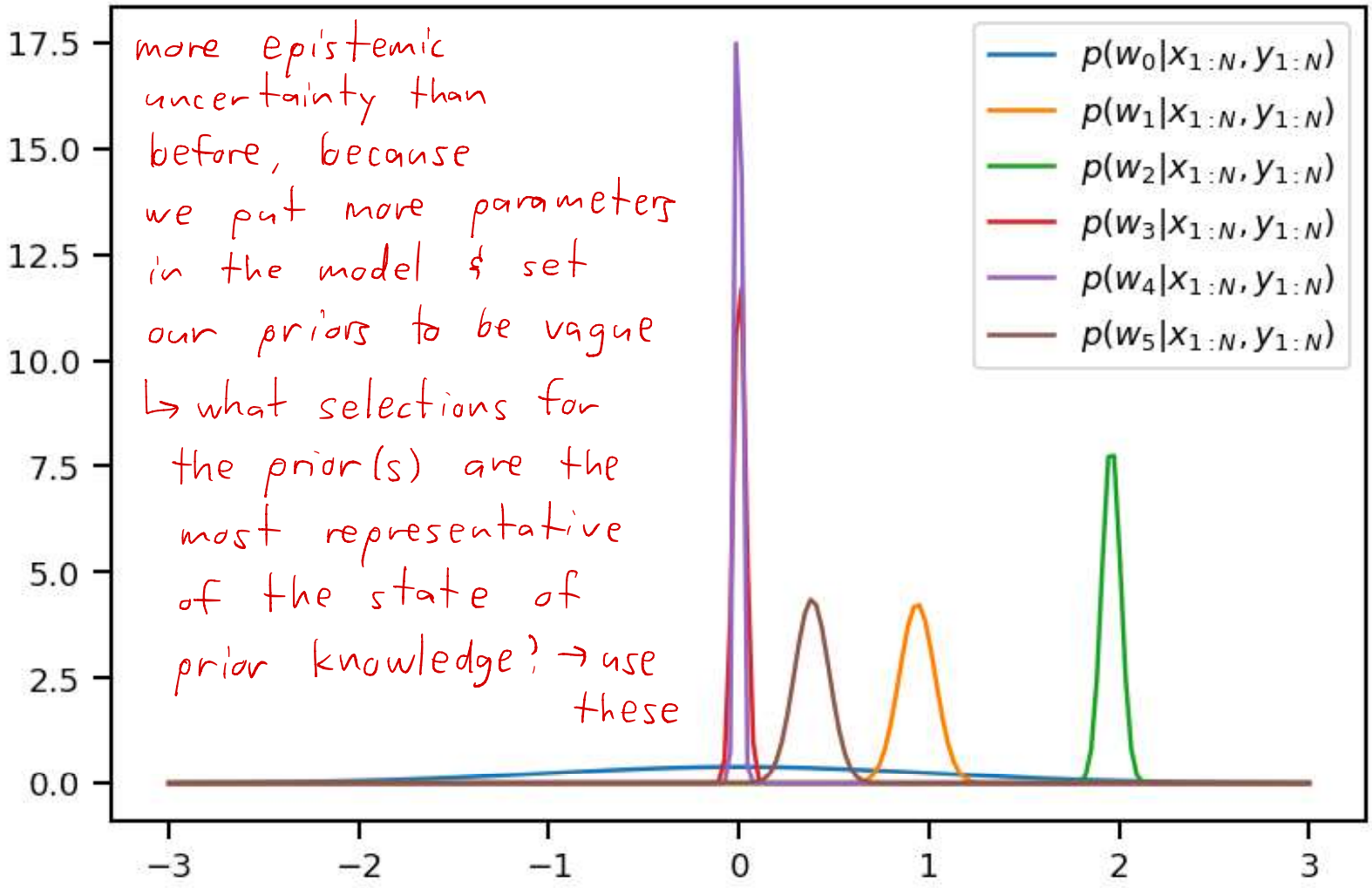
Example



Optimized values for the α_j's



Marginal posteriors for the weights



Open questions

- Cannot be used to compare generalized models with other models (e.g., of completely different functional form). For this, we will need Bayesian model selection.

- How can we model the fact that our noise is input-dependent (heteroscedastic)?