Lecture 15: Advanced topics in Bayesian linear regression

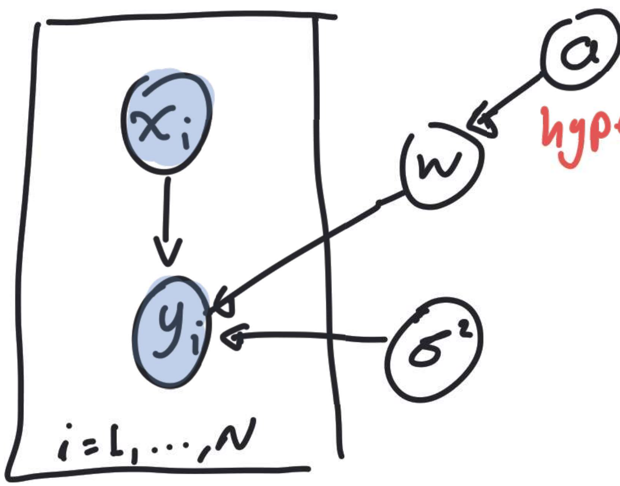
Professor Ilias Bilionis

The evidence approximation

Open questions

- How do I quantify the measurement noise?
- How do we avoid overfitting?
- How do I quantify epistemic uncertainty induced by limited data?
- How do I choose any remaining parameters?
- How do I choose which basis functions to keep?

Hyper-priors



hyper-parameter

Prior: $\alpha \sim p(\alpha)$

$$\underline{w}|\alpha \sim p(\underline{w}|\alpha) = N(\underline{w}|0, \alpha^{-1}\underline{\underline{I}})$$

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

Likelihood: $y_i|x_i, \underline{w}, \sigma^2 \sim N(\underline{\varphi}^T(x_i)\underline{w}, \sigma^2)$

$$p(y_{1:N}|x_{1:N}, \underline{w}, \sigma^2) = \prod_{i=1}^{N} \dots$$

Posterior over hyper-parameters and the evidence approximation

$$p(\underline{w}, \alpha, \sigma \mid x_{1:N}, y_{1:N}) \propto p(y_{1:N} \mid x_{1:N}, \underline{w}, \sigma)\, p(\underline{w} \mid \alpha)\, p(\alpha)\, p(\sigma)$$

$$p(\alpha, \sigma \mid x_{1:N}, y_{1:N}) = \int p(\underline{w}, \alpha, \sigma \mid x_{1:N}, y_{1:N})\, dw$$

$$\propto \int p(y_{1:N} \mid x_{1:N}, \underline{w}, \sigma)\, p(\underline{w} \mid \alpha)\, \boxed{p(\alpha)\, p(\sigma)}\, dw$$

$$= \int p(y_{1:N} \mid x_{1:N}, \underline{w}, \sigma)\, p(\underline{w} \mid \alpha)\, dw \;\; p(\alpha)\, p(\sigma)$$

$$p(\underline{w} \mid x_{1:N}, y_{1:N}, \alpha, \sigma) \propto p(y_{1:N} \mid x_{1:N}, \underline{w}, \sigma) \cdot p(\underline{w} \mid \alpha)$$

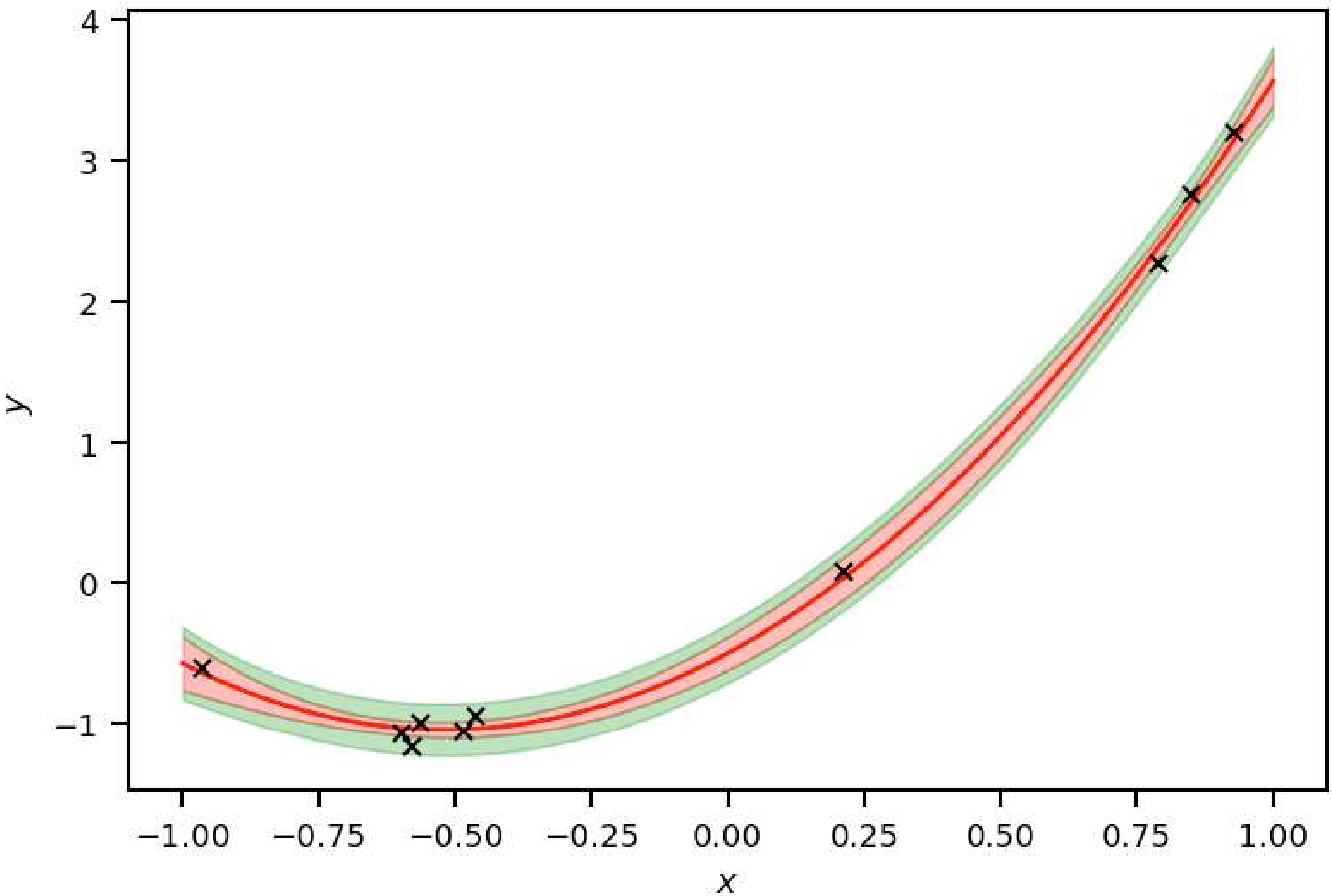
$$\to N(\underline{w} \mid \underline{m}(\alpha, \sigma), \underline{\underline{S}}(\alpha, \sigma))$$

$$= Z(\alpha, \sigma)\, p(\alpha)\, p(\sigma)$$

Evidence approximation : Find α, σ by maximizing their posterior (with w integrated out).

$$\alpha^*, \sigma^* = \arg\max_{\alpha, \sigma} p(\alpha, \sigma \mid x_{1:N}, y_{1:N})$$

Example



Lecture 15: Advanced topics in Bayesian linear regression

Professor Ilias Bilionis

Automatic relevance determination

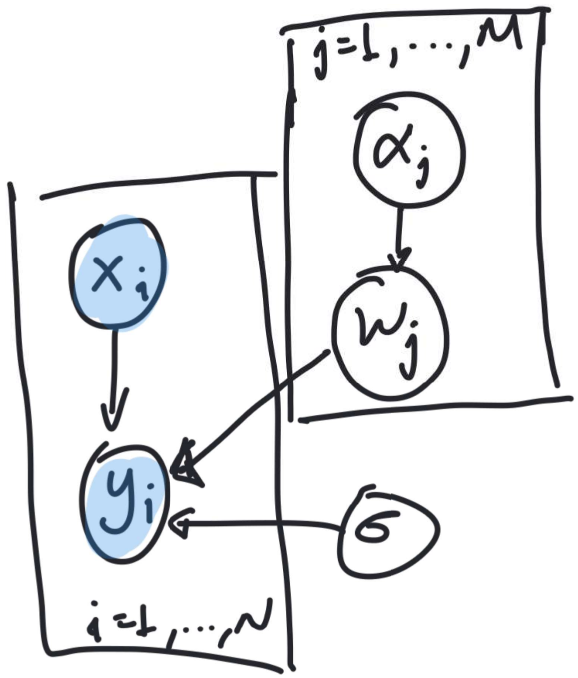
PREDICTIVE
SCIENCE LABORATORY

Open questions

- How do I quantify the measurement noise?
- How do we avoid overfitting?
- How do I quantify epistemic uncertainty induced by limited data?
- How do I choose any remaining parameters?
- How do I choose which basis functions to keep?

PREDICTIVE
SCIENCE LABORATORY

Idea: Different hyper-prior per weight



Prior:

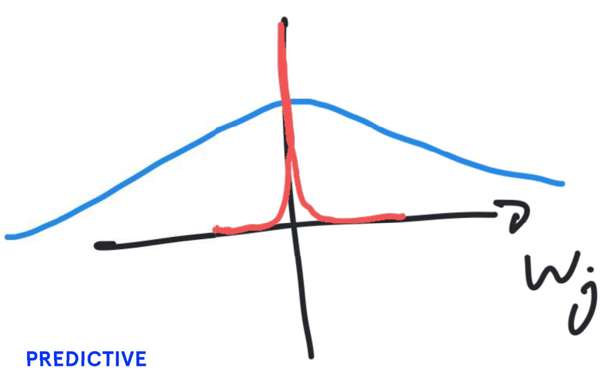
$$\alpha_j \sim p(\alpha_j)$$

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

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

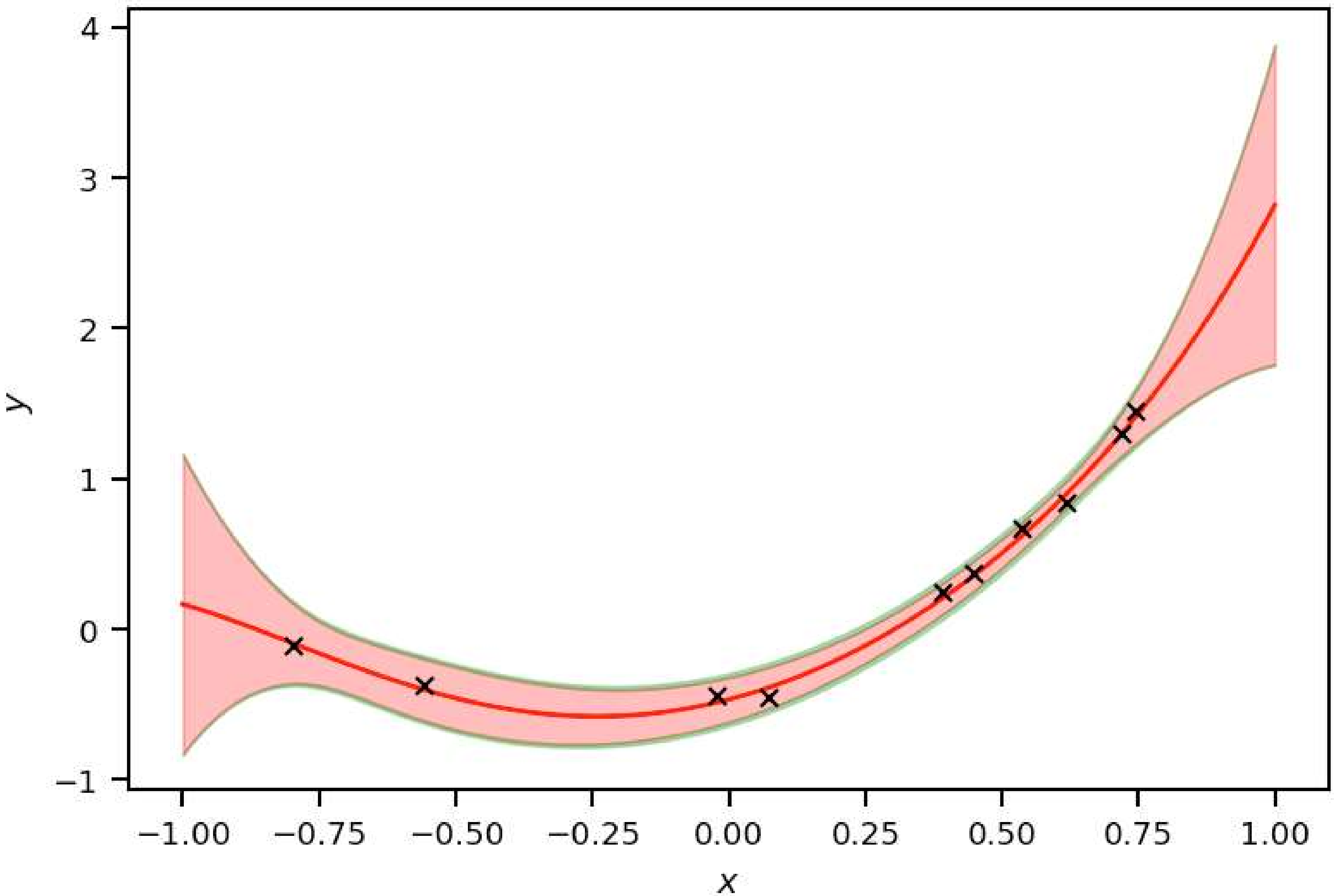
Likelihood:

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



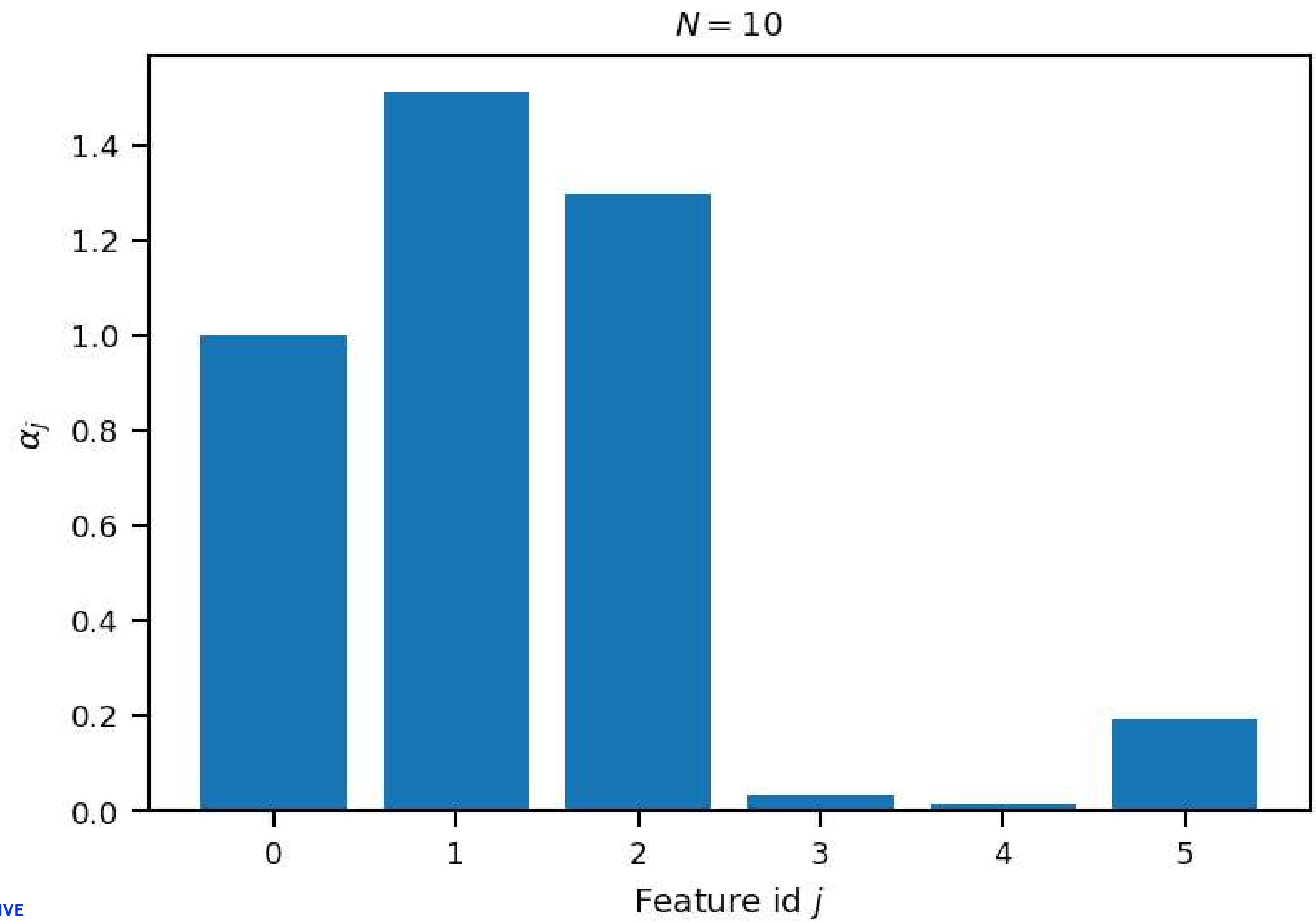
$\alpha_j >> 1$

Example



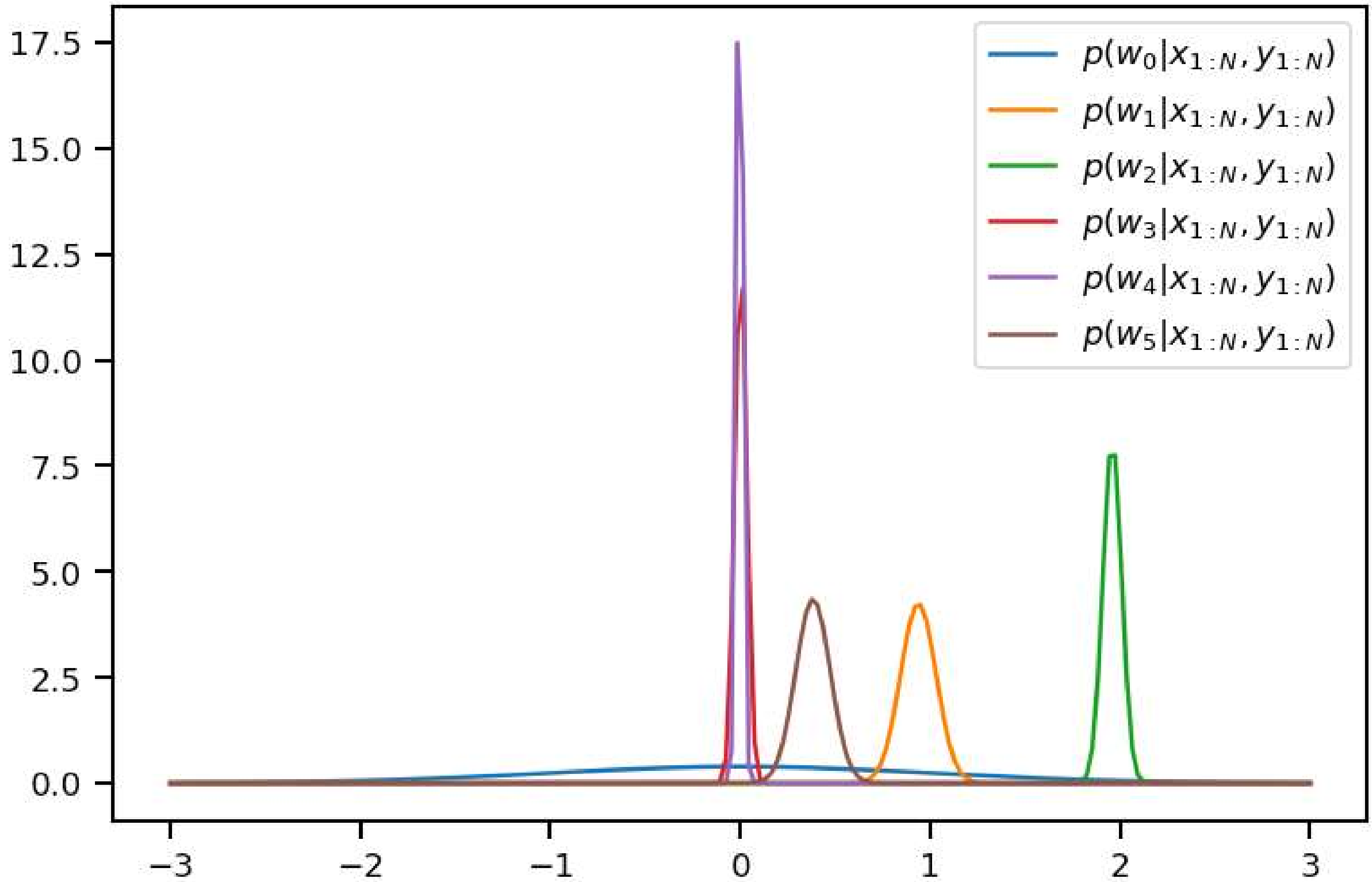
PREDICTIVE
SCIENCE LABORATORY

Optimized values for the α_j's



PREDICTIVE
SCIENCE LABORATORY

Marginal posteriors for the weights



PREDICTIVE
SCIENCE LABORATORY

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)?

PREDICTIVE
SCIENCE LABORATORY

Lecture 15: Advanced topics in Bayesian linear regression

Professor Ilias Bilionis

Diagnostics for posterior predictive

PREDICTIVE
SCIENCE LABORATORY

Standarized errors

Post. Pred. $p(y|x, \text{data}) = N(y|m(x), \sigma^2(x))$

Validatory data : x_i, y_i, $i=1,\dots,N^v$

model says: $y_i | x_i \sim N(m(x_i), \sigma^2(x_i))$

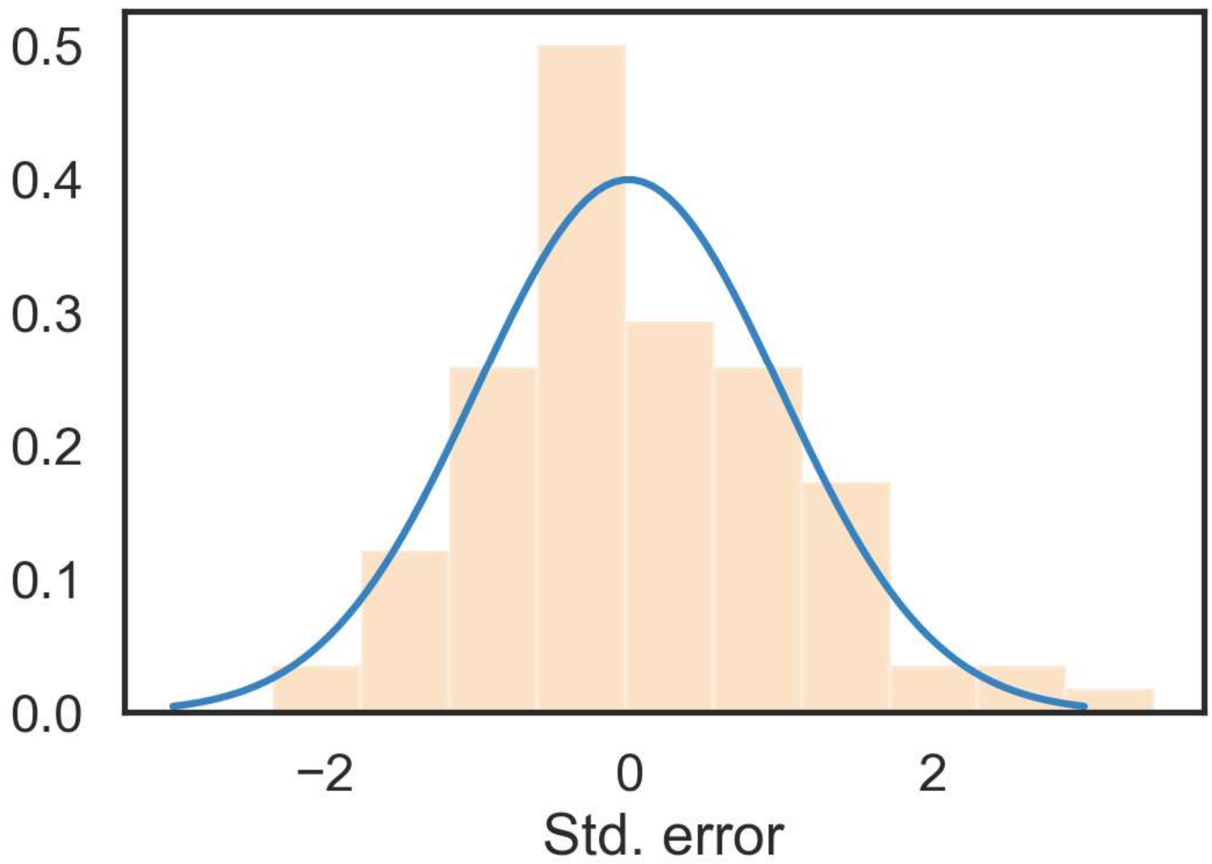
Standarized Error : $z_i := \frac{y_i - m(x_i)}{\sigma(x_i)} \sim N(0,1)$ (If model correct)

$\mathbb{E}[z_i] = \mathbb{E}\left[\frac{y_i - m(x_i)}{\sigma(x_i)}\right] = \frac{\{\cancel{\mathbb{E}[y_i]} - m(x_i)\}}{\sigma(x_i)} = 0$, $\mathbb{E}[y_i] = m(x_i)$

$\mathbb{V}[z_i] = \mathbb{V}\left[\frac{y_i - m(x_i)}{\sigma(x_i)}\right] = \frac{1}{\sigma^2(x_i)}\mathbb{V}[y_i] = \frac{\sigma^2(x_i)}{\sigma^2(x_i)} = 1$

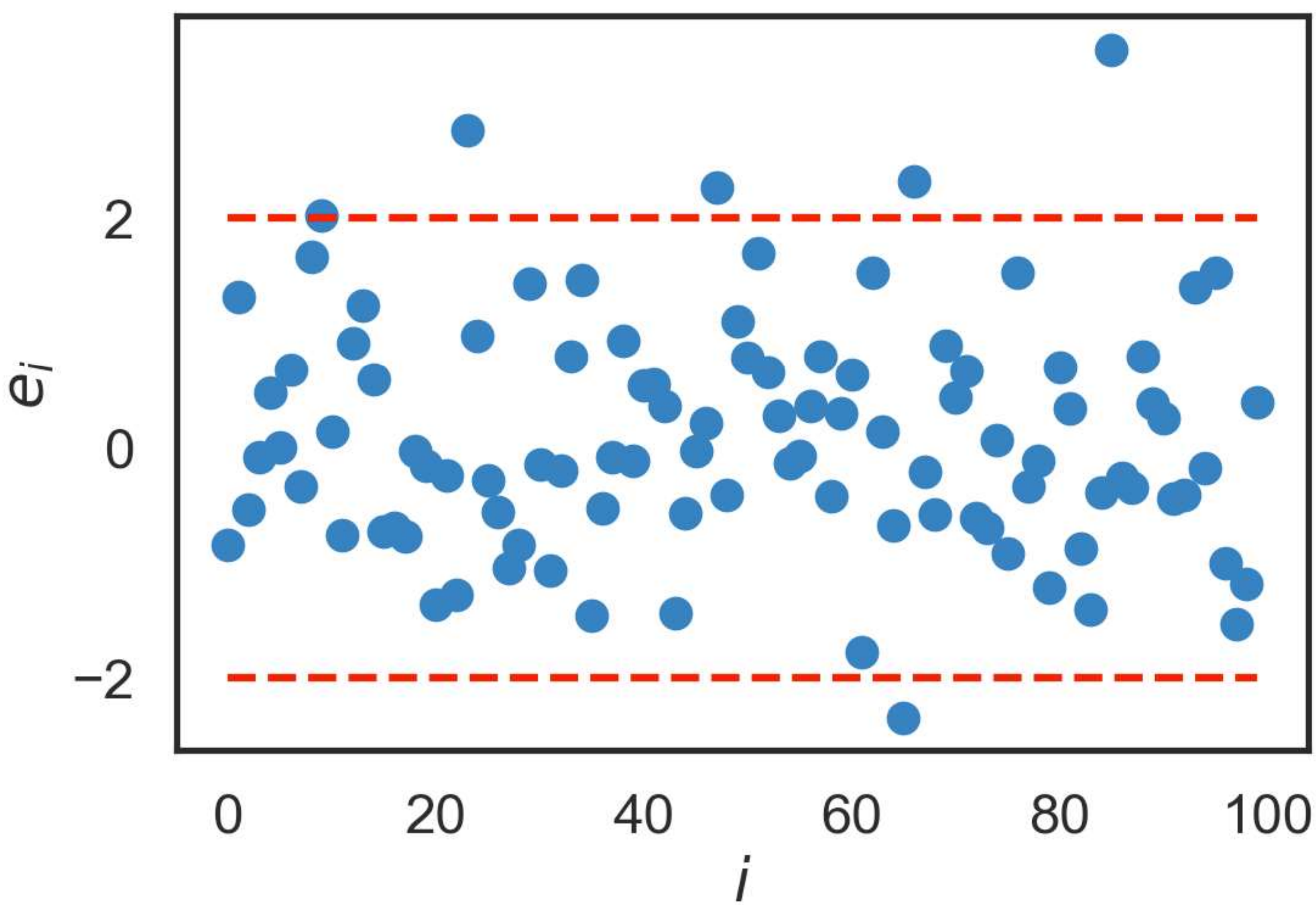
PREDICTIVE
SCIENCE LABORATORY

Standardized Errors



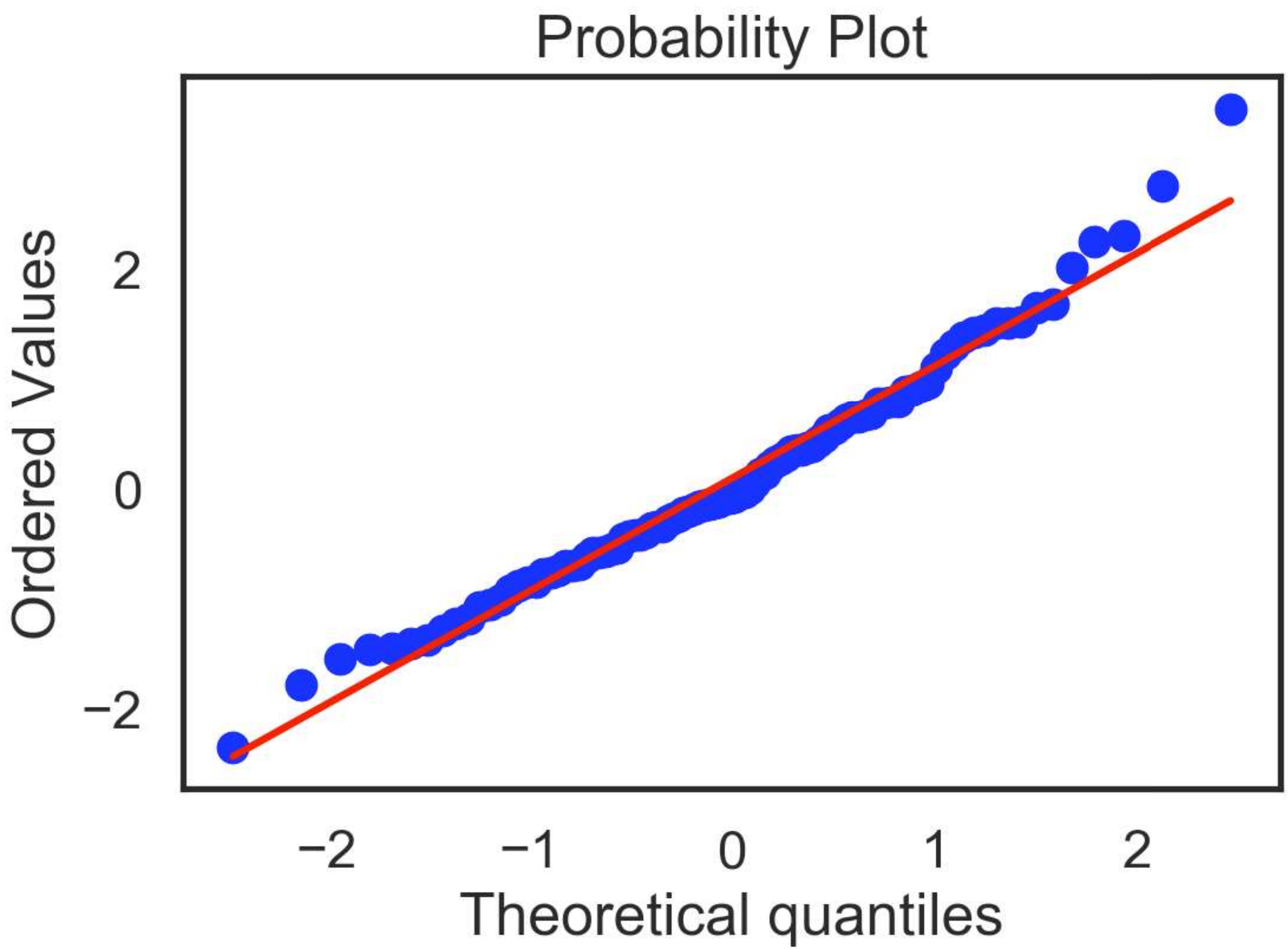
PREDICTIVE
SCIENCE LABORATORY

Standardized Errors



PREDICTIVE
SCIENCE LABORATORY

Standardized Errors



PREDICTIVE
SCIENCE LABORATORY